The Role of DNA Damage Signalling in Adaptive Immunity

Submitted by Rikke Morrish

As a thesis for the degree of

Master of Science by Research in Biological Sciences

In December 2015

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: _____

1 Abstract

Adaptive immunity is essential for the survival of many different organisms. Immune system mechanisms differ between species, but they all have two things in common; adaptability and memory of the immune response. These depend on regulated and often irreversible alterations to the host genome. DNA damage signalling and repair is therefore closely linked with adaptive immunity. In this thesis, I will present the first part of two projects dealing with adaptive immunity of bacterial and mammalian organisms.

I will explore the novel repair factor BRD8 and its potential role in both the DNA damage response pathway and in antibody diversification of mammals. BRD8 was implicated in CSR in a genome-wide short hairpin RNA screen. Preliminary data further implicates BRD8 in both DNA damage repair and antibody diversification. Murine BRD8 is highly expressed in immune cells, specifically immune cells undergoing antibody diversification. Both human and mouse BRD8 share homology with and interact with various known DNA damage repair proteins. Further studies will help reveal its exact role in both DNA damage repair and antibody diversification.

I will also use the bacterial adaptive immune system, Clustered regularly-interspaced short palindromic repeats (CRISPR)-Cas9 for gene editing. Modifications to the system has great potential for therapeutic and experimental uses. A model system based on 293T/GFP-puro cells was developed for rapid characterization of various Cas9 fusion variants we designed.

RNA-deaminase Adar1 has been implicated in somatic hypermutation (SHM) [1], which in antibody diversification increases the affinity of an antibody for a specific antigen through mutations of the variable domain. I therefore fused the deaminase domain of Adar1 to the nuclease-deficient dCas9 in an effort to mimic somatic hypermutation *ex vivo*. The effects of Cas9 was assayed on the DNA level using a celery juice extract we have also developed and established. Preliminary data did not reveal any Adar1 deaminase activity on the eGFP locus, but ongoing studies seem more promising. A functional Adar1-dCas9 can potentially be used for *in vivo* single-base substitutions for both therapeutic and research purposes.

2 Acknowledgements

I would like to thank Dr. Richard Chahwan who provided the projects. Thank you for the opportunity and for support throughout.

I would also like to thank Emily Sheppard and Laurence Higgins for their help and for creating a lovely work environment.

Lastly, thank you to the Danish government for funding my work.

Table of Contents

3 List of Tables

4 List of Figures

5 Introduction

Different immune system mechanisms can be found in different types of organisms, in agreement with their differences and the different types of pathogens they may expect to encounter. In bacteria, the CRISPR system specializes in the RNA mediated DNA cleavage of invading bacteriophage genomes. In vertebrates the system is much more complex, consisting of several specialized cell types. While the two systems differ in both their components and mechanisms, they share two common traits that are essential for their efficiency and speed; *adaptability* and *memory*. The ability of these organisms to store the recognition potential for pathogens and, upon subsequent encounters, utilise this to swiftly react to and dispose of the threat is the essence of the so-called adaptive immune response.

The CRISPR-Cas9 system is a relatively new discovery, but it has already been harnessed as a powerful tool for genomic engineering. The vertebrate immune system, on the other hand, has been studied for more than a century, yet it has proven more elusive in terms of the underlying molecular mechanisms. As of yet, the *ex vivo* mimicry of adaptive immunity has been far less successful than CRISPR. The adaptability of immune systems is a complex topic, which can be approached from several angles. This thesis will tackle two of these angles, with particular emphasis on the overarching themes of adaptive immunity, and its tight association with the DNA damage response. In the first part of the thesis, I will explore a novel repair factor and its potential role in regulating the production and diversification of vertebrate antibodies *in vivo*. In the second part, I will seek to reduce the same process to its bare essentials and utilise the bacterial immune system CRISPR-Cas9 to mimic it *ex vivo*.

5.1 CRISPR-Cas9

The CRISPR-Cas immune systems, which are found in both bacteria and archaea [2], differ slightly from organism to organism. Common among them all are the two core components: the Cas protein capable of cutting DNA, and CRISPR targeting RNA (crRNA) that leads it to specific parts of the foreign DNA.

As CRISPR depends on RNA-based targeting, it is highly specific and easily adaptable. This makes CRISPR advantageous compared to the protein-based

systems, such as zinc-finger nucleases (ZFNs) and transcription-activator-like effector nucleases (TALENs).

While some CRISPR-Cas system types require multiple Cas proteins, Type II requires only one; Cas9. For simplicity's sake, the CRISPR genomic engineering tool is therefore based on this system.

5.1.1 The Type II CRISPR-Cas9 system relies on dual RNA-based targeting

RNA-based targeting of foreign DNA is found in most bacteria and archaea[3]. The specificity of the system relies on the integration of genetic elements from the invading pathogen into a specific region of the bacterial genome. The integrated foreign DNA, so-called protospacers, are interspaced by short palindromic repeats. The whole region is known as the clustered regularly interspaced short palindromic repeats (CRISPR) and has thus given the CRISPR-Cas9 system its name [3].

RNA, transcribed from these DNA regions, is processed and becomes the crRNA that guides the Cas9 protein specifically to the same invading bacteriophage genome if encountered a second time. A transactivating crRNA (tracrRNA) mediates the binding between the crRNA and Cas9. Cas9 cleaves the DNA at the protospacer adjacent motif (PAM). Importantly, the PAM is not contained within the protospacer or

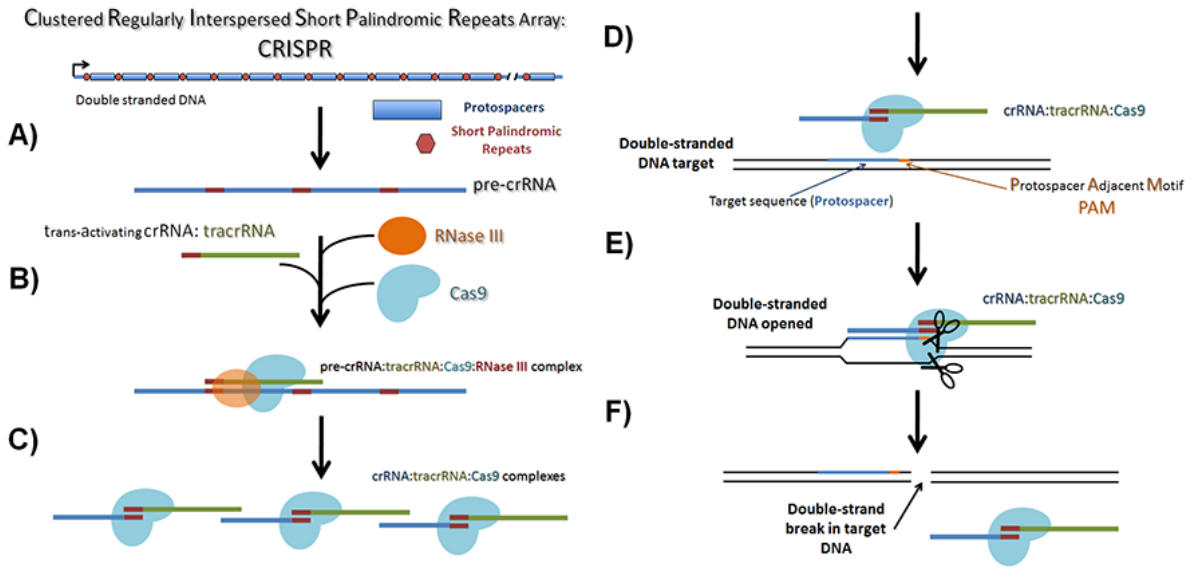

Figure 1: The CRISPR/Cas Type II system[140].
A. The CRISPR region is transcribed into pre-crRNA
B and C. The tracrRNA mediates pre-crRNA processing and Cas9 binding.
D, E and F. The crRNA:tracrRNA:Cas9 binds the target sequence and Cas9 creates a DSB at the PAM site.

crRNA sequence, thus preventing the cleavage of the host CRISPR locus [4,5]. The PAM recognised by Cas9 differs between species [4,6,7].

5.1.2 The CRISPR-Cas9 system can be hijacked and used for genome engineering

The Type II CRISPR-Cas9 system forms the basis of the genome engineering CRISPR tools. It consists of two components; the Cas9 nuclease and a guide RNA (gRNA). The gRNA is a synthetically designed RNA sequence that combines the functions of the crRNA and tracrRNA [8] (see section 5.1.1). Cas9 has been codon optimised for expression in several species, including humans. The Church and Zhang labs produce the most commonly used vectors for expressing Cas9 and gRNA in several organisms.

Designing suitable gRNAs for a given target sequence has one main criterion: the presence of the PAM sequence immediately downstream of the gRNA complementary sequence. Cas9 from *Streptococcus pyogenes* (*SpCas9*) is the most commonly used Cas9. The PAM sequence recognised by the SpCas9 is NGG [9], making gRNA design possible for almost any region of a given genome in most organisms. Several bioinformatics tools have been developed for gRNA design. These consider potential off target sites, and ease the choice of the best suited gRNA for a given target [10] (http://www.addgene.org/crispr/reference/#grna).

The original CRISPR-Cas9 system was used for creating knock-outs. The DSB created by the Cas9 is generally repaired through the non-homologous end joining (NHEJ) pathway, which often results in a deletion or insertion [8]. Statistically, two thirds of the time (66%) this process causes a frameshift, and thus a knockout of the gene.

Cas9 binding to DNA through interaction with gRNA, and its ability to cleave DNA, are two independent functions. A dCas9 mutant unable to cleave DNA can therefore still be targeted to a specific region via gRNA. This allows for a multitude of modified CRISPR-Cas9 systems to be created, expanding the possible applications beyond gene knock-outs. Several proteins have been fused to dCas9, including other nucleases [11], fluorophores [12] and transcriptional activators and repressors [13,14].

5.2 The vertebrate immune system

Unlike bacteria, vertebrate pathogenicity relies mostly on protein, lipids, and sugar moieties. That is why the adaptive immune system of jawed vertebrates relies on antibodies, or immunoglobulins (Ig), for the recognition of invading antigens [15]. Mutations and recombination of the DNA at the Ig locus allows the immune system to produce a vast repertoire of antibodies from a single locus. This intentional DNA damage takes place in two developmental stages of B cells and, unsurprisingly, is tightly regulated [16,17]. Without these constraints, an organism would suffer from immune deficiencies and be susceptible to pathogen invasion, while increased or uncontrolled DNA damage can cause accumulation of mutations elsewhere in the genome and lead to cancerous growth [18].

5.2.1 Antibody diversification

The primary antibody diversification process of V(D)J recombination occurs in the bone marrow. It involves the rearrangement of the variable (V), joining (J) and diverse (D) segments to make up the variable region of the Ig gene. At this stage, the B cells are capable of producing a wide array of antibodies. However, their affinity for antigens remains limited and, as they maintain expression of the same precursory effector region, their potential effector functions are restrained. Effector functions determine the mechanism of pathogen elimination, as well as antibody localisation [19].

Upon antigen engagement, B cells enter germinal centres in lymphatic tissue where secondary antibody diversification takes place. Somatic hypermutation (SHM) of the variable region increases the affinity of the antibody for specific antigens, while class-switch recombination (CSR) exchanges the IgM constant region with other constant regions such as IgG and IgA. Both of these processes are initiated by the small mutator protein activation-induced cytidine deaminase (AID). AID deaminates cytosines resulting in U:G mismatches in the DNA sequence.

For SHM, AID mutations initiate the accumulation of point mutations in the variable region, altering the tertiary structure of the antibody and thus allowing increased affinity for encountered antigens. In CSR, AID exerts its activity at switch regions upstream of the different constant regions. Here the AID-initiated damage leads to double-strand breaks (DSBs), allowing CSR to occur (see Fig.3B-E).

5.2.2 SHM involves mutations at both G-C and A-T base pairs

During SHM the mutation rate at the variable region of the Ig locus increases to 10^{-3} mutations per base pair per generation. This is 10^6 times higher than the background mutation rates in the cells [20,21]. All four bases are mutated, with C:G and A:T pairs being targeted approximately equally [21,22]. All mutations depend on the activity of AID [23], but only the C:G pairs are directly altered by AID. The mutations affecting A:T pairs are dependent on the mismatch repair system; in particular Msh2 and Msh6 [24–26]. The Msh2-Msh6 heterodimer has been proposed to recognise the AID-generated U:G mismatches and recruit low-fidelity polymerases which amplify the mismatches in the variable region [26,27].

5.2.2.1 AID initiates SHM by deaminating cytosines

AID deaminates cytosine residues to create U:G mismatches. It has been shown that *in vitro* AID acts only on ssDNA, and not RNA, dsDNA or DNA/RNA hybrids [28]. It preferentially targets RGYW/WRCY (R = A/G, W=A/T and Y=C/T) hotspot motifs [22]. C and G mutations occur at a similar frequency, indicating that AID targets both the transcribed strand (TS) and non-transcribed strand equally [29].

The U:G lesions can result in a mutated DNA sequence by way of different mechanisms through either replication or erroneous repair. First, as dU is read as dT by the replication machinery, one daughter cell would acquire a G-to-A mutation during replication during cell division (should the U:G mismatch go unnoticed). Second, if the dU is excised, error-prone base excision repair can then potentially fill the gap with any base. Third, the aforementioned mismatch repair can recognise the mismatch and through non-canonical pathways recruit error-prone polymerases; opening the sequence up for a myriad of mutations at both C:G and A:T pairs [21]. One such error-prone polymerase is polymerase eta (Pol η).

5.2.2.2 Adenine mutations cannot be explained solely by DNA polymerase η activity

Pol η has been shown to be important for the A:T mutations of SHM [30]. Pol η is a translesion synthesis polymerase. It functions to allow the replication machinery to

bypass UV-induced damage by accurately replicating DNA with thymine-dimers. However, its fidelity when replicating non-damaged bases is low, with base-substitutions occurring at a rate of 10^{-2} to 10^{-3} [31,32]. Patients with xeroderma pigmentosum have mutations in Pol η and are sensitive to sunlight. Sequencing data from such patients reveal that in the absence of Pol η, the Ig variable regions have a decrease in A:T mutations without affecting the overall mutation rate [33,34]. Similar patterns have been observed in mice deficient in Pol η [35–37]. Furthermore, Pol η has been shown to have a preference for WA motifs when mis-incorporating bases; WA being the A:T hotspot in somatic hypermutation [38].

While this all appears to be fairly conclusive, issues have been raised with this interpretation of the Ig variable region sequencing data [39,40]. The main focus of these critiques is that while C and G mutations caused by AID occur with a similar frequency on both strands, the A and T mutations exhibit a strand bias with a 2-fold increase in mutations of A relative to T on the non-transcribed strand [41]. Although reports have shown that Pol η contributes to this strand bias [42], a plausible mechanism behind this bias has yet to be illustrated [39,40]. Why would Pol η be recruited to the non-transcribed strand more than the transcribed strand, when AID targets both equally? The mechanism is not known, but AID expression levels [29] and the UNG and Msh2/Msh6 pathways [43] are contributing factors. It appears that there is still a long way to go before we resolve the full mechanism behind somatic hypermutation. It certainly opens up the possibility of additional factors contributing to the generation of the A:T mutations.

5.2.2.3 The reverse transcriptase model of SHM

One theory, put forward by Edward J. Steele, involves an RNA intermediate. It is argued that SHM stems from both DNA and RNA mutations which are coupled by a reverse transcriptase [39]. Edward J. Steele and colleagues have shown that Pol η itself can function as a reverse transcriptase *in vitro* [30]. This theory allows the possibility of the involvement of RNA editing proteins in SHM, including the Adenosine Deaminase Acting on RNA-1 (Adar1), which is responsible for A-to-I editing in dsRNA [44].

Steele proposes that Adar1 deaminates adenosines in dsRNA stem loops of nascent mRNA [39] based on a statistical correlation between the frequency of WA-to-WG mutations and the number of predicted mRNA hairpins for the site [45]. Pol η

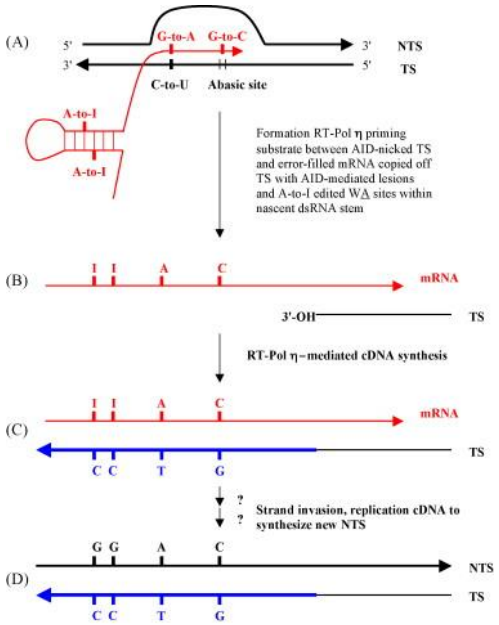

Figure 2: Edward J. Steele's 'Reverse transcriptase model of SHM[39]*: transcription-coupled DNA and RNA deamination and reverse transcription'*

would then reverse transcribe the mutated mRNA into cDNA, which would serve as a template for replicating the non-transcribed strand [39] (see Fig.2).

5.2.2.4 Adar1 deaminates adenosines near mismatches

Steele's reverse transcriptase model for SHM is interesting and strongly supported by robust statistical data. However, there is very little supporting experimental evidence for this complex theory. Therefore, a different scenario should also be considered: the possibility that Adar1 also exerts its adenosine deaminating activity on DNA. Indeed, two isoforms of Adar1 exist; the constitutively expressed shorter isoform, and the interferon-inducible longer isoform which contains an N-terminal Z-DNA binding domain [46–48]. The biological role of this Z-DNA binding domain has yet to be clarified, but it has been proposed that it directs Adar1 to highly transcribed genes [49]. Z-DNA is a left-handed double helix [50]. Negative supercoiling, which appears upstream of an active RNA polymerase, has a stabilizing effect on Z-DNA [51,52].

Recently, a correlation between overexpression of Adar1 and an accumulation of somatic DNA mutations in Bcl6-deficient cells was shown [1]. Bcl6 is a transcriptional repressor [53], which inhibits cell senescence [54]. It has also been implicated in antibody diversification; Bcl6 knockout mice have impaired germinal centre B cell formation [55–57]. Sequencing of the Ig-Sμ region in Bcl6 deficient IgG1 cells revealed an increase in somatic mutations. The majority of the mutations were A:T mutations. Adar1 was found to be overexpressed in the Bcl6 deficient cells. Furthermore, exogenous overexpression of Adar1 can induce accumulation of somatic hypermutation without AID expression [1].

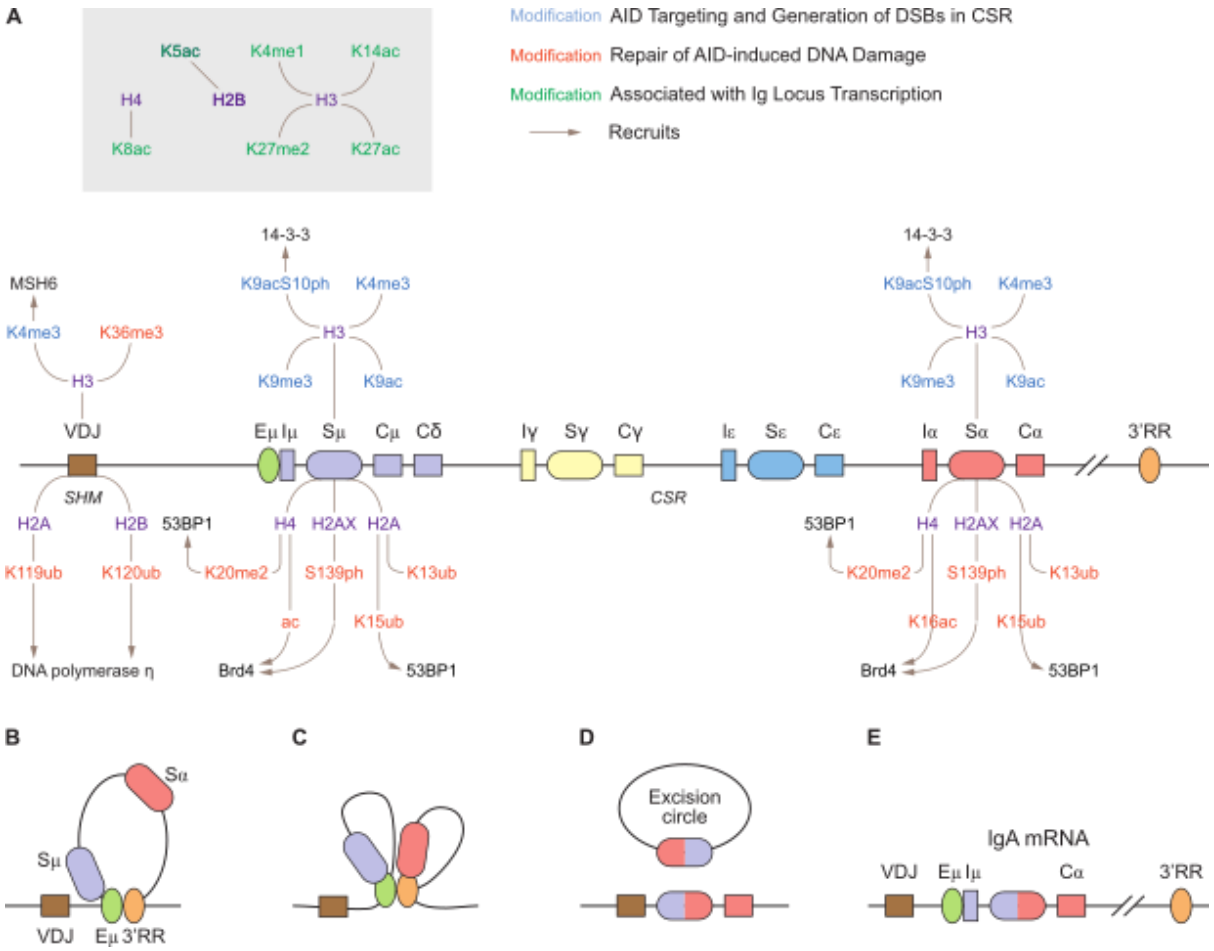

Figure 3: Schematic illustration of chromatin modifications at the Ig locus during secondary antibody diversification.*
A. The Ig locus and the histone modifications that contribute to the regulation of SHM and CSR. Green modifications are associated with Ig locus transcription. Blue modifications are integral to AID targeting and generation of DSBs in CSR. Orange modifications are involved in signalling and protein recruitment during the repair phase subsequent to AID-induced DNA damage.
B and C. The Eμ enhancer associates with the 3' regulatory region (3'RR), bringing switch regions into close proximity.
D. Class-switch recombination occurs when the DSBs are repaired in such a way that the DNA segment between the two switch regions is lost.
E. The Ig locus after CSR; now expressing the IgA antibody.
** The figure was made in collaboration with Emily Sheppard and Helen Jones.*

One quality of Adar1 that is especially interesting, when seen in the light of a possible role in SHM, is its preference for deaminating adenosines near mismatches. The secondary structure of the targeted dsRNA has been shown to determine the specificity and efficiency of Adar1 adenosine deamination [58,59]. The specificity or efficiency of Adar1 deamination is not dependent on its dsRNA binding domains[58]. The efficiency of deamination is decreased in Adar1 constructs with mutated Z-DNA binding domains, but only when editing minimal substrates; 15 bp dsRNA stem [58]. Which adenosine within a sequence Adar1 targets is dependent on mismatches within the region. A mismatched adenosine is preferentially edited by Adar1[58]. Editing can also occur in the absence of this mismatch, but for those substrates a G:U mismatch occurred nearby [58]. Recently, it has been shown that the requirement for a mismatch at the targeted adenosine is dependent on the surrounding bases with the 5' neighbour having the biggest influence. The data also support the notion that a mismatch at a nearby C:G pair increases specificity and efficiency in the absence of a mismatch at the targeted adenosine [59].

I postulate that, if Adar1 is indeed capable of deaminating deoxy-adenosines, the mismatches induced by AID may serve as a substrate recruitment platform. Furthermore, as the Ig locus is highly transcribed [60], Z-DNA formation is likely and could also help recruit Adar1 to the region. More studies are needed in order to elucidate the possible role of Adar1 in somatic hypermutation.

5.2.3 Successful CSR depends on chromatin modifications

Both SHM and CSR are orchestrated by chromatin modifications. Recruitment of AID, and the repair of AID induced DNA damage, are in large part regulated by signalling cascades involving histone modifications; including ubiquitination, phosphorylation, methylation and acetylation (see Fig.3). The following introductory chapters will focus on the chromatin modifications that are essential for successful CSR to occur. It is worth noting that additional regulatory and signalling elements are involved, including non-coding RNA [61].

5.2.3.1 Donor and recipient switch regions are marked by histone modifications prior to DSBs

Specific stimuli drive CSR in B cells towards a particular Ig isotype. The exact mechanism with which the donor and the desired recipient switch region are selectively targeted, has not been fully elucidated, but specific histone modifications have been implicated as markers for these regions prior to DSB initiation (see Fig.3A). It seems plausible that these histone marks could function as recruitment platforms for AID or other components of the CSR machinery.

H3K4me3 has been implicated as a marker for IgA switching in CH12F3-2A cell lines[62,63]. It is facilitated by the FACT complex, a potent chromatin modifying complex. Knockdown of the FACT components SSRP1 and SPT16 caused a general decrease in H3 methylation in the Sμ region and a specific reduction of the H3K4me3 mark in the Sα region, correlating with a significant decrease in IgA switching[63]. The reduction in H3K4me3 led to an decrease in DNA breaks at the switch regions, however the intermediate steps between histone mark and DSB creation remain to be elucidated [62,63]. Knockdown of the histone chaperone Spt6 also leads to a decrease in the H3K4me3 mark in Sμ and Sα, along with a subsequent reduction in DSBs in the two regions[64]. This further supports the role of H3K4me3 in CSR and implicates Spt6 as a regulator of the mark. Interestingly, the Spt6 knockdown also led to a decrease in H3K4me3 in the variable region, losing SHM [65], while previous reports did not see any effect in the variable region [66].

The histone modifications H3K9me3 and H3K9ac are also found at S regions involved in recombination[67]. They depend on cytokine stimulation but not AID expression. Therefore, they are thought to proceed, or maybe even be directly involved in, AID recruitment to the switch regions and subsequent recombination [67]. Outside the Ig locus the H3K9me3 mark has been shown to play a role in DSB repair rather the generation of DSBs. It functions through a direct interaction with the histone acetyltransferase Kat5, and abrogation of the H3K9me3 mark causes defective DSB repair [68,69] (see section 5.2.3.3). This implies that H3K9me3 may have a role in both targeting and repair of DSBs in CSR.

A fourth combinatorial histone mark of H3K9ac and H3S10ph, has also been implicated as a marker of donor and recipient S regions in CSR[70]. The Sμ donor

region was found to be enriched in H3K9acS10ph. Upon cytokine activation, the desired recipient S region also became enriched in the mark [70]. A direct interaction between H3S10ph and the 14-3-3 adapters has been determined. The addition of an acetyl group to K9 of the same histone increases the affinity of this interaction [71,72]. This is particularly interesting because the 14-3-3 complex binds AID and 5'-AGCT-3' motifs that are found in both switch regions and the variable region. Furthermore, upregulation of 14-3-3 correlates with increased CSR [73]. Altogether this hints at a role for H3K9ac/H3S10ph as a recruiter of 14-3-3 to mediate the recruitment of AID to the switch regions.

5.2.3.2 Histone modifications recruit DSB repair proteins in CSR

Post AID-induced DSB creation, two things are important. First, the DSBs must be repaired to maintain genome stability. Second, the DSBs must be repaired in such a way that the donor and desired recipient switch regions are recombined and the intervening DNA is lost (see Fig.3D). The repair proteins needed for this process are recruited through a signalling cascade that includes multiple histone modifications. One protein that has a well-established role in DSB repair, including CSR-dependent DSBs, is 53BP1 [74–80]. Various pathways involving chromatin modifications are involved in the recruitment of 53BP1 to DSB sites.

The histone mark H4K20me2 recruits 53BP1 through a direct interaction with the Tudor domain of 53BP1 [81,82]. Lack of H4K20me2 enrichment at DSBs, caused by the loss of the methyltransferase MMSET, has been shown to result in poor CSR [83].

Some of the earliest measured histone modifications to appear at DSBs are phosphorylated H2AX (γH2AX) and H4ac. The kinase ATM is responsible for phosphorylating H2AX in DSBs. In typical DNA repair, it has been shown that the activation of ATM prior to H2AX phosphorylation is dependent on the aforementioned Kat5. Upon binding to H3K9me3 (see section 5.2.3.1), through its chromodomain, Kat5 recruits and acetylates ATM, which in turn autophosphorylates [84]. 53BP1 binds γH2AX through its BRCT domain and aides DSB repair by bridging the broken ends [85]. Kat5 is also responsible for H4 acetylation [69].

The BET family member Brd4 interacts with acetylated histones through its two bromodomains [86] and has been found to play a role in DSB repair during CSR [87]. Indeed, knockdown or inhibition of Brd4 resulted in a decrease in CSR in CH12F3-2A cells (REF). It also reduced 53BP1 and UNG (another DNA repair protein) occupancy at switch regions. Importantly it did both without affecting the H4 acetylation. Chromatin Immunoprecipitation (ChIP) and immunoprecipitation assays confirmed an interaction between Brd4 and 53BP1, H4ac and γH2AX. On this basis, Brd4 has been proposed to function as a chromatin-bound platform for 53BP1 and UNG [87].

Finally, a ubiquitination pathway has also been implicated in the recruitment of 53BP1 to DSBs [88]. It involves two E3 ubiquitin ligases; RNF8 and RNF168. RNF8 is initially recruited to DSBs by the ATM mediated phosphorylation of MDC1 (mediator of DNA damage checkpoint 1) at a TQxF consensus site. Upon recruitment, RNF8 ubiquitinates H2A histones at the DSB site. This in turn recruits RNF168 which amplifies the ubiquitin signal through polyubiquitination. Knockdown of either RNF8 or RNF168 causes reduction in the accumulation of 53BP1 at AID-induced DSBs as measured by immunofluorescence in CH12F3-2A cells[88]. A concomitant reduction in CSR was also measured [88]. In recent studies, RNF168 has been shown to monoubiquitinate H2A on lysine 13 and 15 [89,90]. In accordance with this, 53BP1 has been shown to bind H2AK15ub directly through an ubiquitination-dependent recruitment motif [82]. So it seems the ubiquitin mark itself mediates the recruitment of 53BP1. Indeed, the reduced accumulation of 53BP1 at DSBs in RNF8 and RNF168 deficient cells can be rescued with the expression of an ubiquitin-H2AX fusion protein [91].

It is apparent that a myriad of proteins are involved in CSR. However, there are still large gaps in our understanding of the “epigenetic” mechanisms that target AID to regulate efficient CSR and antibody diversification in general. This opens up the possibility of novel proteins serving important roles. A genome-wide short hairpin RNA screen our lab has previously conducted has implicated, among other hits, the bromodomain protein BRD8 in CSR and possibly other DNA damage response pathways too.

5.2.3.3 BRD8 is a chromatin reader

Relatively little is known about BRD8 protein. From its sequence we can infer that it contains one, possibly two, bromodomains (see section 7.2.1 and 7.2.2) which are known to bind acetylated protein moieties [92]. This is particularly exciting in light of the multitude of chromatin modifications (including acetylations) that are essential for CSR. As mentioned previously, the related protein Brd4 has been shown to have a role in CSR (see section 5.2.3.2). Furthermore, it has been reported that human bromodomain-containing proteins play an essential role in the DNA damage response [93].

Overexpression of BRD8 has been implicated as an important factor in colorectal cancer[94]. Indeed, it was linked to the increased resistance to chemotherapy of these cancers[94]. The resistance is caused by reduced spindle poison sensitivity. Spindle poisons activate the spindle checkpoint, initiating mitotic arrest followed by cell death or senescence. BRD8 siRNA mediated knockdown in various cell lines showed increased sensitivity to spindle poisons [94]. This further implicates BRD8 in the DNA damage response and associates it with adequate maintenance of genomic stability.

BRD8 is sometimes known as TRCp120, because it was originally identified as a nuclear receptor coactivator of the thyroid hormone receptor [95]. Its potential role in transcription activation has later been supported by immunoprecipitation assays revealing that BRD8 is a subunit in the NuA4 HAT complex [96,97], which functions in transcriptional activation [98]. More importantly, when considering a possible role for BRD8 in CSR; the NuA4 HAT complex, of which BRD8 is a member, also has a role in DSB repair [98]. In fact, the previously mentioned Kat5, also known as Tip60, is the catalytic subunit of the NuA4 HAT complex.

5.3 Aims and objectives

5.3.1 Developing a modified CRISPR-Cas9 system

The CRISPR-Cas9 genome engineering system has been modified before in various ways. Fusing the inactive dCas9 to other proteins allows us to utilise the readily adaptable RNA-directed targeting and combine it with the catalytic activity of another protein. In this project, I will construct an Adar1-dCas9 fusion protein to allow specific targeting of Adar1 dependent point mutations. To do this, I must:

1. Design and construct an Adar1-dCas9 fusion expression system
2. Design suitable gRNAs for targeting wildtype Cas9 and Adar-Cas9 to eGFP
3. Assay the effects of Cas9 variants using celery juice extract assay and flow cytometry

5.3.2 Assessing the role of BRD8 in antibody diversification and DNA damage repair

BRD8 has been implicated as a possible important factor in both the DNA damage response and antibody diversification. In this project I will begin to unravel the role of BRD8 in these interconnected processes. I will:

1. Assess the likelihood of a role for BRD8 in the DNA damage response and in antibody diversification using available data
2. Transfect 3T3 cells with BRD8 gRNAs
3. Develop methods for verifying deletions in 3T3 and CH12 cells

6 Methods and Materials

6.1 Developing a modified CRISPR-Cas9 system

6.1.1 Constructing Adar-Cas9 expression plasmid

The human codon-optimized *Streptococcus pyogenes* Cas9 nuclease with mutations D10A and H840A (Addgene #47314 [99]) was used as expression vector. The plasmid is hereby referred to as the dCas9 plasmid.

Primers were designed to amplify the Adar1 deaminase domain from FLIS-Adar1 plasmid, kindly provided by Joana M. P. Desterro, while adding a SpeI restriction site, a FLAG-tag and a XTEN-linker at the C-terminus of the deaminase domain. In addition, a XbaI restriction site was included at the N-terminus. The final PCR amplicon results in a 1391 bp segment, using primers RCOL289 and RCOL290, respectively:

5'-
ACTAGTACTAGTATGGACTACAAGGACGACGATGACAAACCCAAGAAGAAGAGGAAGGTGAAGGCAGAACGCATGGGTTTC-3'

5'-
TCTAGATCTAGAGCACTTTCGGGTGTGGCGGACTCTGAGGTCCCGGGAGTCTCGCTGCCGCTTACTGGGCAGAGATAAAA-3'.

Both the dCAs9 plasmid and the FLIS-Adar1 plasmid contains an ampicillin resistance gene. The Adar insert was therefore amplified through two consecutive PCRs in order to avoid contamination with the FLIS-Adar1 plasmid in the ligation colonies. The first PCR was performed with RedTaq® ReadyMix™ PCR Reaction Mix (R2523, Sigma Aldrich) in 25µl with 100µM of each primer, 1 ng plasmid DNA and a program of: 94˚C x 5min, 20 cycles of (94˚C x 30s, 55˚C x 30s, 72˚C x 1min10s), 72˚C x 5min. The second PCR was performed in a similar 25µl reaction with 1 µl of a 1/500 dilution of PCR#1 as DNA template and a program of: 94˚C x 5min, 30 cycles of (94˚C x 30s, 55˚C x 30s, 72˚C x 1min10s), 72˚C x 5min.

The Adar insert was digested with FastDigest restrictions enzymes SpeI (BcuI) and XbaI, and the m2 plasmid with FastDigest restriction enzyme XbaI (FD0684, Thermo Scientific), for 7min at 37˚C in FastDigest buffer. 330 ng Adar insert was

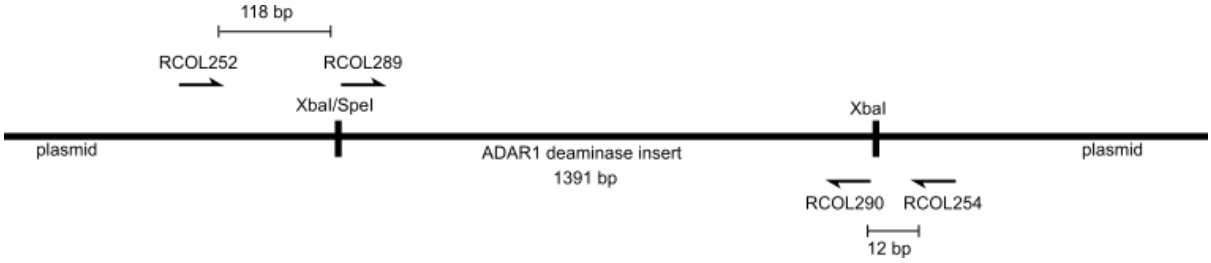

Figure 4: Genetic map for the ligation of Adar1 deaminase insert into m2 plasmid. Restriction sites and primer locations are marked.

ligated into 90 ng linearized m2 plasmid in a 10 µl volume using Rapid DNA Ligation Kit (K1422, Thermo Scientific). The ligation mix was left at room temperature for 10 min. The ligated plasmid was transformed into NEB 5-alpha Competent *E.coli*; 5-10µl DNA added to 50µl competent cells, then incubated on ice for 30 min, heat shocked at 42˚C for 45s, on ice for 5 min and left to recover in LB at 37˚C, and plated on LB, 100µg/ml ampicillin plates. Plates were left at 37˚C overnight.

The following day colonies were picked, dipped in PCR mixture for colony PCR and used to inoculate overnight cultures in 5ml LB, 100µg/ml ampicillin media. Colony PCR was performed with RedTaq® ReadyMix™ PCR Reaction Mix (R2523, Sigma Aldrich) in 12.5µl with 2µM of each primer and a program of: 94˚C x 5min, 30 cycles of (94˚C x 30s, 55˚C x 30s, 72˚C x 1min10s), 72˚C x 5min. The primers used were RCOL252 (5'-CGCAAATGGGCGGTAGGCGTG-3') and RCOL254 (5'-TGGAGTACTTCTTGTCCATGGTGGCA-3'). They are situated in the plasmid sequence on either side of the XbaI restriction site (see Fig.4).

Overnight cultures testing positive for the insert were spun down and plasmid was purified from the cell pellets using GeneJET Plasmid Miniprep Kit (K0502, Thermo Scientific). Concentrations of plasmid DNA was measured by Nanodrop. The orientation of the insert in the plasmids was confirmed by subsequent PCR reactions performed using RedTaq® ReadyMix™ PCR Reaction Mix (R2523, Sigma Aldrich) in 25µl with a program of: 94˚C x 5min, 30 cycles of (94˚C x 30s, 55˚C x 30s, 72˚C x 1min10s), 72˚C x 5min, 10ng of plasmid and 2µM of each primer for the primer combinations RCOL252/RCOL290 and RCOL252/RCOL289 (see Fig.4). Positive

results were verified by sequencing (Eurofins) using RCOL252 and RCOL321 (5'-CGATCGGTATTGCCCAGAAC-3').

6.1.2 EGFP gRNA design and plasmid construction

In order to test the activity of the Adar-Cas9 fusion protein compared with other Cas9 variants, gRNAs targeting eGFP of the 293T/GFP-puro cell line (AKR-202, 2BScientific) were designed. Suitable gRNAs were chosen based on scores given by the gRNA prediction softwares CRISPRdirect [10] and Feng Zhang lab's Target Finder (http://crispr.mit.edu/), as well as the expected activity of the Adar-Cas9 fusion protein.

The designed gRNAs were synthesized according to the Church hCRISPR gRNA synthesis protocol (https://www.addgene.org/crispr/church/) and ligated into the gRNA_Cloning Vector (Addgene #41824[100]).

6.1.3 Transfection of 293T/GFP-puro cells

All 293T/GFP-puro cell growth took place in DMEM media, 2µg/ml puro (DMEM-puro media) at 37˚C. Using polyethylenimine (PEI), cells were transfected with a combination of the constructed eGFP gRNA and Adar-Cas9 plasmids, the original FLIS-Adar1 plasmid, wtCas9 plasmid (Addgene #41815[101]), m2Cas9 plasmid (Addgene #47314[99]) and myc gRNA plasmid previously constructed by Laurence Higgins. Two dish/well sizes were used and the cell number and DNA concentrations were adjusted accordingly. Below are the methods for the 6-well plate. Cells transfected in 6cm dishes were incubated with triple the amount of DNA. In accordance with this, the amount of cells seeded was also tripled.

For the 6-well plate ~100,000 293T/GFP-puro cells/well were seeded and 2 ml DMEM-puro media was added the day before transfection. The DNA mix was prepared in 250µl OPTIMEM/well in Eppendorf tubes. A total of 5µg of DNA, 2.5µg Cas9 variant plasmid and 2.5µg gRNA plasmid, was incubated with 15µl PEI (3µl PEI per 1µg DNA) at room temperature for 15 min. For control samples containing only Cas9 or gRNA plasmid the total DNA amount was 2.5µg and the PEI volume 7.5µl. Post incubation DNA mix was added to each well and the cells were left for 3 hours before the PEI was removed through a media change.

6.1.4 Celery Juice Extract Assay

Analysis of the effects of Cas9 at the DNA level was based around a mismatch-specific nuclease. The assay consisted of four steps; 1. PCR and PCR purification, 2. Heteroduplex formation, 3. Digest with mismatch-specific nuclease and 4. Visualisation of the reaction on a gel. The mismatch-specific nuclease was extracted from celery.

6.1.4.1 Preparation of celery juice extract

The celery juice extraction was modified from a published protocol [102]. One bunch of celery was washed and blended with water. The blended celery was filtered and spun down for 20 min at 2600 g. The supernatant was transferred to a clean tube. 1 M Tris and phenylmethanesulfonyl fluoride (PMSF) was added to the solution for a final concentration of 0.1 M Tris-HCl, pH 7.7, 100 µM PMSF.

144 g $(NH_4)_2SO_4$ was added per litre of solution for a final solution of 25% $(NH_4)_2SO_4$. The solution was mixed for at 4˚C for 30 min and spun down 20,000 g at 4˚C for 30 min. The supernatant was transferred to a clean tube.

390 g $(NH_4)_2SO_4$ was added per litre of solution for a final solution of 80% $(NH_4)_2SO_4$. The solution was mixed for at 4˚C for 30 min and spun down 20,000 g at 4˚C for 60 min. The supernatant was discarded. The pellet was resuspended in 0.1M Tris-HCl pH 7.7, 0.5M KCl, 100 µM PMSF followed by overnight dialysis in 10L of same buffer.

6.1.4.2 eGFP primer design

Primers were designed for the eGFP DNA sequence using Primer3 [103,104]. Primers with suitable melting temperatures and location in the sequence were chosen.

6.1.4.3 Verifying the CJE assay with control DNA

Two variants of control DNA with known mismatches between them were amplified from two plasmids and purified using the GeneJET PCR Purification Kit (K0701, Thermo Scientific) by Angelina Block.

The purified DNA samples were normalized to a concentration of ~50ng/µl. Two 20 µl samples in Phusion HF buffer (F-518L, Thermo Scientific) were prepared for heteroduplex formation; control DNA variant 1 alone and a 1:1 ratio of the two variants. To facilitate heteroduplex formation, DNA was denatured and reannealed in thermocycler under the following conditions: 95˚C x 10 min, 95-85˚C at -2˚C s^{-1}, 85 ˚C x 1 min, 85-75˚C at -0.3˚C s^{-1}, 75˚C x 1 min, 75-65˚C at -0.3˚C s^{-1}, 65˚C x 1 min, 65-55˚C at -0.3˚C s^{-1}, 55˚C x 1 min, 55-45˚C at -0.3˚C s^{-1}, 45˚C x 1 min, 45-35˚C at -0.3˚C s^{-1}, 35˚C x 1 min, 35-25˚C at -0.3˚C s^{-1}, 25˚C x 1 min, 25-10˚C at -0.3˚C s^{-1}, 10˚C hold.

For the digest 4µl of CJE was added, and the solution was left to incubate for 45 min at 45˚C. Afterwards, 2µl 0.25M EDTA was used to stop the reaction, and samples were run with loading dye on a 1% agarose TAE buffer gel at 110V for 30 min. The gel was stained with Midori Green Advance (MG04, Geneflow) and visualized under UV.

6.1.4.4 Analysis of the Cas9 effects on the GFP locus

Genomic DNA was extracted using the GeneJET Genomic DNA Purification Kit (K0721, Thermo Scientific) from both the transfected cells and untransfected control cells. The DNA was eluted in 100µl 55˚C water. Using suitable primers (see section 6.1.4.2) the region expected to contain the mismatch was amplified using PCR performed with RedTaq® ReadyMix™ PCR Reaction Mix (R2523, Sigma Aldrich) in 50µl with 100-200ng genomic DNA, 2µM of each primer and a program of: 94˚C x 5min, 30 cycles of (94˚C x 30s, 55˚C x 30s, 72˚C x 30s), 72˚C x 5min. The PCR products were purified using the GeneJET PCR Purification Kit (K0701, Thermo Scientific).

The purified DNA samples were normalized to a concentration of ~50ng/µl. Three 20 µl samples in Phusion HF buffer (F-518L, Thermo Scientific) were prepared for heteroduplex formation; wild-type DNA, experimental DNA and a 1:1 ratio of the

two. The experimental DNA was then treated in the same manner as the control DNA (see section 6.1.4.3).

6.1.5 Flow cytometry

Cells were fixed in 1% PFA/PBS. Using a BD FACSCalibur the GFP signal of GFP negative 293A cells, and transfected and control 293T/GFP-puro cells was measured and analysed.

6.2 Assessing the role of BRD8 in antibody diversification and DNA damage repair

6.2.1 BRD8 sequence conservation and ancestry

BRD8 protein sequences from different species were retrieved from NCBI Reference Sequence Protein database [105]. Sequences from *Homo sapiens, Pan troglodytes, Mus musculus, Felis catus, Pteropus vampyrus, Chelonia mydas, Xenopus (Silurana) tropicalis, Danio rerio, Xiphophorus maculatus, Esox lucius, Stegodyphus mimosarum, Drosophila melanogaster* and *Ustilago maydis 521* were aligned in Jalview [106] using Muscle multiple alignment. The long isoforms were also aligned separately using the same conditions. Domains were predicted using the SMART software [107].

6.2.2 Interaction networks of human and mouse BRD8

The interaction networks of human and murine BRD8 were examined using the STRING database [108] of known and predicted protein interactions. The information is based on different lines of evidence originating from four different sources; genomic context, high-throughput experiments, conserved co-expression and previous knowledge. On the basis of these lines of evidence, each protein-protein interaction is given a confidence score.

6.2.3 Expression levels of BRD8

Expression levels of BRD8 in different tissues and cell types was assessed based on RNA seq data from BioGPS [109] and microarray data from Immgen [110]. The BioGPS RNA seq raw data was processed and normalized using MEF values by Leo Bennett.

6.2.4 Subcellular localization of BRD8

Subcellular localization data for BRD8 in U-2OS cells was available from Protein Atlas [111].

6.2.5 BRD8 gRNA design

Two gRNAs for exon 2 of the BRD8 gene in *M. musculus* were designed by Leo Bennett and ligated into px260-U6-DR-BB-DR-Cbh-NLS-hSpCas9-NLS-H1-shorttracr-PGK-puro vector (Addgene #42229[4]).

6.2.6 3T3 cells transfection

All 3T3 cell growth took place DMEM media at 37˚C. The day before transfection ~100,000 3T3 cells/well were seeded in a 6-well plate and 2 ml DMEM media was added. On the day of transfection the DNA mix was prepared in 250µl OPTIMEM/well in eppendorf tubes. A total of 5µg of DNA was incubated with 15µl PEI (3µl PEI per 1µg DNA) at room temperature for 15 min. Post incubation DNA mix was added to each well and the cells were left for 3 hours before the PEI was removed through a media change. Genomic DNA was extracted after 72 hours (see section 6.2.8).

6.2.7 CH12 cells

CH12 cells previously transfected with the BRD8 gRNAs and frozen by Leo Bennett were thawed, let recover and pelleted for genomic DNA extraction (see section 6.2.8).

6.2.8 Celery Juice Extract Assay

Genomic DNA was extracted using the GeneJET Genomic DNA Purification Kit ((K0721, Thermo Scientific) from both the transfected cells and untransfected CH12 and 3T3 cells. The DNA was eluted in 100µl 55˚C water.

Using primers RCOL270 (5'-CTACTATCGTGTGACTGATCCTGG-3') and RCOL271 (5'-CTTAAATTGGAGGGATGGCTGTAC-3') a 668bp region around the gRNA targets was amplified using PCR performed with RedTaq® ReadyMix™ PCR Reaction Mix (R2523, Sigma Aldrich) in 50µl with 100-200ng genomic DNA, 2µM of each primer and a program of: 94˚C x 5min, 30 cycles of (94˚C x 30s, 55˚C x 30s, 72˚C x 35s), 72˚C x 5min. The PCR products were purified using the GeneJET PCR Purification Kit (K0701, Thermo Scientific) with elution in 50µl 55˚C water.

The purified DNA samples were normalized to a concentration of ~50ng/µl. Three 20 µl samples in Phusion HF buffer (F-518L, Thermo Scientific) were prepared for heteroduplex formation; wild-type DNA, experimental DNA and a 1:1 ratio of the two. The experimental DNA was then treated in the same manner as the control DNA (see section 6.1.4.3).

6.2.9 Visualising BRD8 localisation using BRD8 antibody

293A human cells and 3T3 murine cells were seeded on 13mm objective glasses in 24-well plates and grown in 0.5ml DMEM media per well. Prior to fixation cells were treated with 1µM camptothecin for 45 min.

Cells were washed in PBS, permeabilized with 2%PFS in PBS for 15 min and blocked with 5% FCS, 0.05% TritonX in PBS. Cells were incubated overnight at 4˚C with Rabbit anti-BRD8 (A300-219A-T, BETHYL) at a 1:100 dilution as per manufacturer's instruction and Ms mAb to γH2AX (ab26350, abcam) at a 1:500 dilution, both in 5% FCS, 0.05% TritonX in PBS. After three washes with 0.05% TritonX in PBS, the cells were incubated at room temperature in the dark for 45 min with secondary antibody conjugates Goat pAb to Ms IgG, Alexa Fluor™ 488 (ab150117, abcam) and Dnk pAb to Rb IgG, Alexa Fluor™ 647 (ab150067, abcam).

7 Results

7.1 Developing a modified CRISPR-Cas9 system

7.1.1 Constructing Adar-Cas9 expression plasmid

The colony PCR revealed that colonies 7, 10, 13, 15, 18 and 19 had taken up plasmid containing the insert, while others had taken up re-ligated m2 plasmid (see Fig.5A). Multiple colonies appeared to contain several plasmid variants. The cause of this can most likely be found in either the transformation step or the picking of the colonies for the PCR. During the transformation single bacterial cells may have taken up more than one plasmid. When preparing the colony PCR, although care was taken, it is possible that two or more colonies were picked as one.

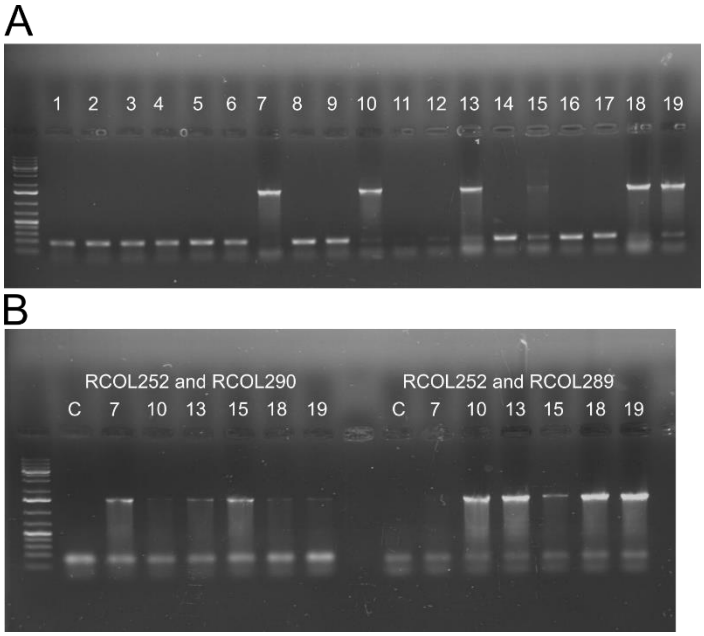

Figure 5: Gels for ligation of Adar1 deaminase insert into m2 plasmid.
A. Ligation colony PCR with primers RCOL252 and RCOL254.
B. Plasmid PCR with primer combinations RCOL252/RCOL290 and RCOL252/RCOL289.

Subsequent PCRs on the purified plasmid from the colonies containing the plasmid with the insert confirmed the orientation of the insert. Colony 7 was the only unambiguous colony with the insert in the correct orientation. Sequencing confirmed the orientation of the insert. The two primers either end of the insert enabled the verification of the entire sequence of the insert and the two ligation sites. Colony 7 was used for subsequent experiments.

7.1.2 eGFP gRNA design

For the design of suitable gRNAs for eGFP, several criteria had to be met; some of which were based on the expected activity of the Adar-Cas9 fusion protein.

First, target options were considered taking two things into account; the 5' (T>A>C>G) and 3' (G>C=A>T) nearest neighbour preferences of Adar1 [112] and potential editing effects of Adar1 on the different amino acid codons. The A to G

mutation caused by Adar1's deaminase activity is not capable of creating a novel stop codon in a sequence (see Table 1), and as such causing a substitution mutation in a mutation sensitive part of the protein sequence was the only option. Therefore, the search for target options was focused on XAG/XAC/XAA sites in the chromophore domain of the eGFP sequence. Tyr-66 (TAC) provided two targeting options both leading to a change in fluorescence. An A to G mutation on the coding strand would lead to a Tyr to Cys substitution (TAC to TGC), which would inactivate the chromophore [113]. Meanwhile, an A to G mutation on the noncoding strand would result in a Tyr to His substitution (GTAG → GTGG = TAC → CAC), leading to a strong cobalt blue signal but only dim fluorescence [114,115].

Table 1: The deaminase activity of Adar1 cannot create novel stop codons in a sequence regardless of which strand is targeted.

A to G mutation on the coding strand	
Creating TAG	Start with TAA, which is already a stop codon
Creating TGA	Start with TAA, which is already a stop codon
A to G mutation on the noncoding strand, causing a T to C mutation on the coding strand	
No Cs in any of the stop codons	

Second, the positioning of the gRNA relative to the chosen target site was considered. Adar1 targets dsDNA [1], and the XTEN-linker should enable the Adar1 deaminase domain to target an adenine base 7-13 base pairs downstream of the

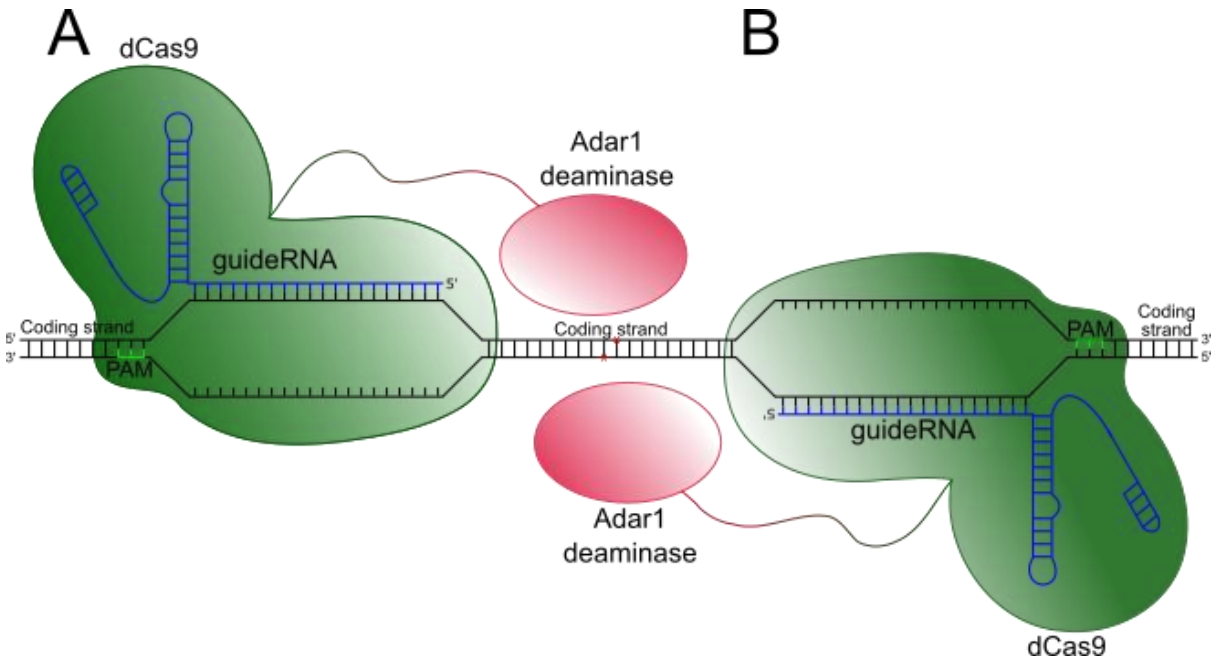

Figure 6: Two options for gRNA design

gRNA without interference from dCas9A [11] (see Fig.6), so this was deemed to be the optimal setup. gRNAs on both coding (Fig.6A) and non-coding (Fig.6B) strand were considered, as either should allow the Adar1 deaminase domain to target Tyr-66.

Table 2: eGFP gRNA stats

eGFP gRNAs		CRISPRDirect			Feng Zhang lab's Target Finder		
	Sequence including PAM	Score	Off target sites	Mismatches between gRNA and offtarget site	20mer	12mer	8mer
R1	CCCTCGTGACCACCCTGACCTAC	86	119 (20 in genes)	2-4	0	6	989
R2	CCCGTGCCCTGGCCCACCCTCGT	46	360 (73 in genes)	2-4	0	74	9041

Third, taking all previous criteria into account, suitable gRNAs were assessed based on their prediction software scores. gRNAs with higher scores were expected to exhibit higher specificity for the selected sequence and thus fewer off-target effects. Two suitable gRNAs were designed; R1, **CCC**TCGTGACCACCCTGACCTAC, and R2, **CCC**GTGCCCTGGCCCACCCTCGT, which were both a compromise between ideal software score (see Table 2) and ideal location. The Tyr-66 codon is found at the last three bases of R1, while it is located 13 bases downstream of R2.

A further two gRNAs targeting the eGFP gene were designed by Laurence Higgins using slightly different criteria.

7.1.3 eGFP primer design

The designed primers, GAGGAGCTGTTCACCGGG and CTTGTACAGCTCGTCCATGC, were expected to amplify a 702 base pair region encompassing all gRNA target sites. The melting temperatures for the primers were 59.73 and 59.00 respectively.

7.1.4 Evaluation of the CJE assay

The activity of the CJE assay was evaluated using control DNA (see Fig.7). The control DNA will create one mismatch upon heteroduplex formation in the A/L sample. Lane 2 containing 'A' DNA on its own, and thus containing no mismatches, shows only the

original uncut band. In lane 3, a further two clear bands appear, indicating digest of the heteroduplexes formed. The full size band is in turn slightly faded.

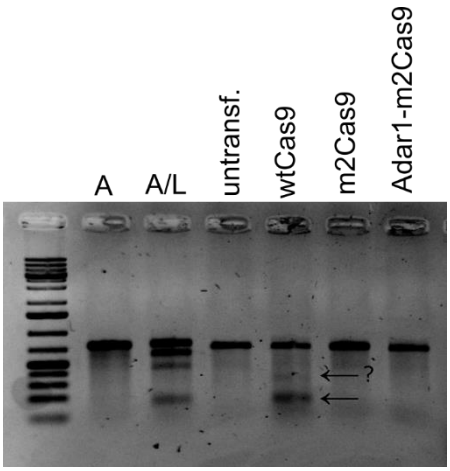

Figure 7: CJE assay of control DNA (lane 2 and 3) and DNA from 293T/GFP-puro cells (lane 4 to 7).
Lane 2 shows the 'A' DNA on its own. No mismatches are expected and the CJE does not cut. Lane 3 shows the mixed A/L DNA. Besides the full sized band, a further three smaller bands have appeared indicating digest at mismatches by the CJE.
Lanes 4 to 7 show amplified region of DNA from 293T/GFP-puro cells. 'untransf.' indicates DNA from untransfected cells. The other lanes are named according to the Cas9 variant they were transfected with. They were all transfected with four eGFP gRNAs. A clear band is visible 4. Digest at mismatches is apparent in lane 5 (wtCas9). The CJE does not reveal mismatches for m2Cas9 or Adar1-m2Cas9.

7.1.5 Assessment of gRNAs and wild-type Cas9 plasmids

293T/GFP-puro cells express eGFP and this fluorescence can be measured by flow cytometry. The signal is distinct from non-fluorescent 293A cells (see Fig.8A). Transfection with eGFP gRNA without any Cas9 variant did not alter the measured fluorescence (see Fig.8B). For all 293T/GFP-puro cell populations transfected with the eGFP gRNAs and wild-type Cas9 I saw a consistent reduction of around 50% in GFP signal as measured by flow cytometry (see Fig.8D).

Transfection with wild-type Cas9 without gRNA did not alter the measured eGFP signal. Transfection with wild-type Cas9 and a myc gRNA exhibited a small change in eGFP fluorescence for a subset of the cells (see Fig.8C).

In Figure 9, three different wild-type Cas9 expressing plasmids were tested with the eGFP gRNAs; one Church plasmid and two Zhang plasmids. The Church plasmid is the expression vector used for the Adar1-Cas9 construct, while the Zhang plasmids are used for the BRD8 deletions (see section 6.2.5). All three had a deletion efficiency, as measured by loss of GFP signal, around 50% (see Fig.S1).

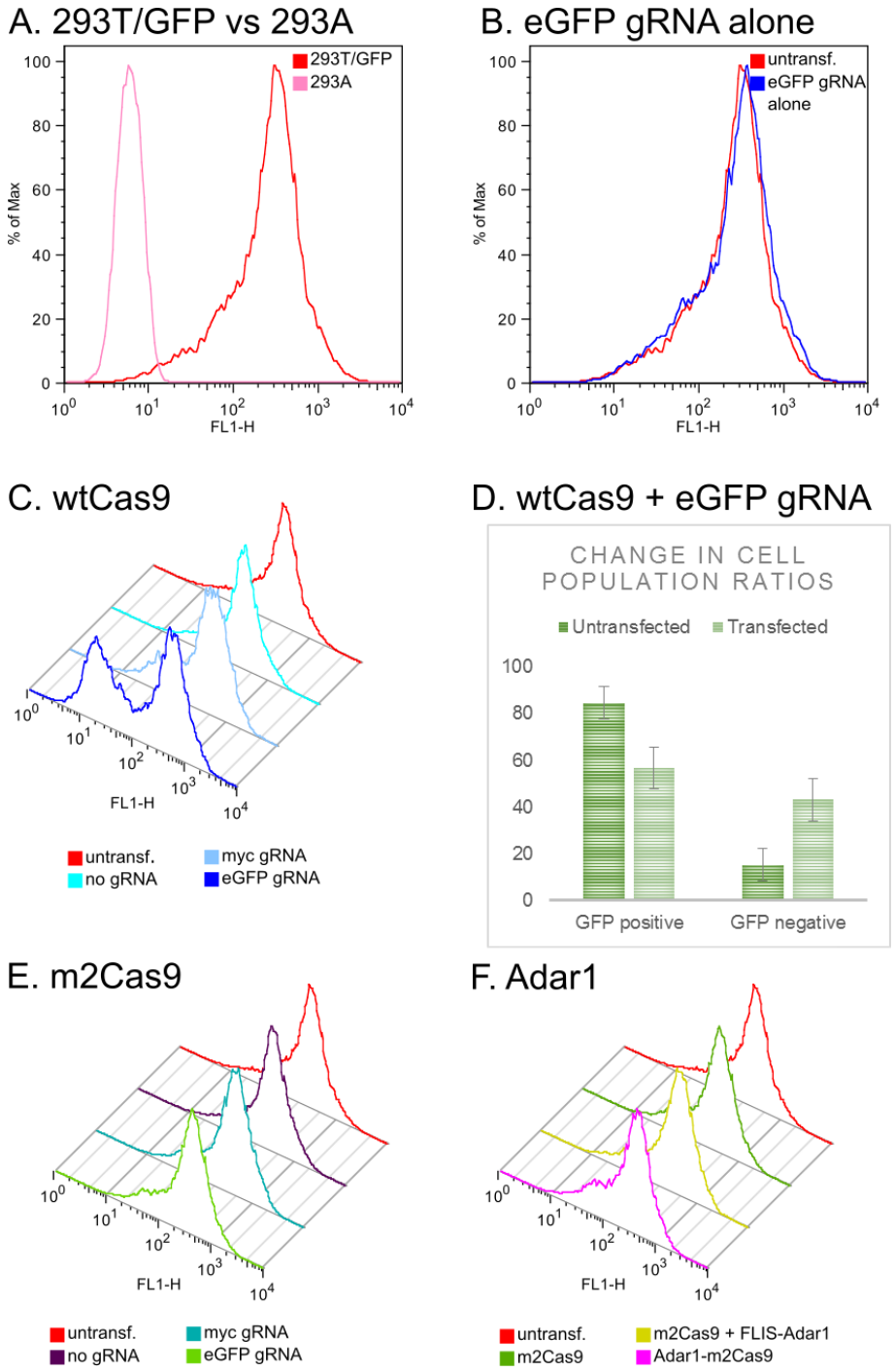

Figure 8: eGFP emission measured by flow cytometry for 293T/GFP-puro cells and 293A cells.
A. eGFP-expressing 293T/GFP-puro cells and non-fluorescent 293A cells.
B. Untransfected 293T/GFP-puro cells and 293T/GFP-puro cells transfected with eGFP gRNA but no Cas9 variant.
C. 293T/GFP-puro cells transfected with wild type Cas9 and no gRNA, myc gRNA or eGFP gRNA.
D. Cells transfected with wtCas9 and eGFP gRNA compared to untransfected cells. The average percentage of GFP positive and negative cell populations based on six biological replicates. Error-bars showing standard deviation.
E. 293T/GFP-puro cells transfected with m2Cas9 and no gRNA, myc gRNA or eGFP gRNA.
F. 293T/GFP-puro cells transfected with eGFP gRNA and m2Cas9, Adar1-m2Cas9 or m2Cas9 and FLIS-Adar1.

The wild-type Cas9 mediated deletion in the eGFP gene was also assessed using the CJE assay (see Fig.7). As seen on the gel, the DNA from the transfected

cells formed mismatches during heteroduplex formation and was cut by the mismatch specific nuclease indicating that the DNA was altered in a subset of the cells. This is consistent with the flow cytometry data.

7.1.6 Assessment of the Adar-Cas9 construct

93T/GFP-puro cells transfected with m2Cas9 and eGFP gRNAs exhibit a small shift in measured fluorescence, seen as a small bump in the left side of the fluorescence peak (see Fig.8E). 293T/GFP-puro cells transfected without gRNA or with myc gRNA do not exhibit this change.

293T/GFP-puro cells transfected with the Adar1-m2CAs9 construct or Adar1 and m2Cas9 separately are compared to 293T/GFP-puro cells transfected with m2Cas9 in Figure 8F. The measured fluorescence does not vary significantly between the samples. The CJE assay did not reveal any mismatches during heteroduplex formation (see Fig.7). This indicates a lack of mutation in the eGFP gene.

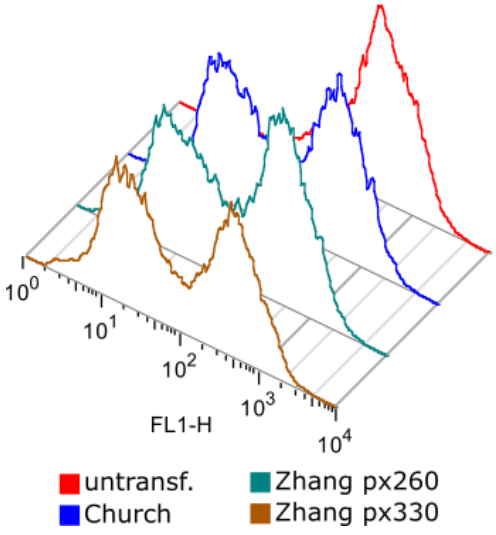

Figure 9: Flow cytometry results for wild-type Cas9 transfections of 293T/GFP-puro cells. One Church plasmid and two Zhang plasmids were tested and compared to an untransfected control.

7.2 Assessing the role of BRD8 in antibody diversification and DNA damage repair

7.2.1 The human BRD8 has several potential interaction sites

Homo sapiens BRD8 has been reported to have at least two isoforms [116,117]; a short isoform of 951 amino acid residues containing one bromodomain and a longer isoform of 1235 amino acid residues containing two bromodomains. In addition to the bromodomains, the predicted domains of both isoforms (visualised in Fig.S2) include coiled-coil and low complexity regions. Both of these regions are mediators of protein-protein interactions [118,119].

7.2.2 BRD8 is conserved through vertebrates

BRD8 is found in several vertebrate species as well as *Drosophila melanogaster*, but for the majority, only a short isoform containing one bromodomain has been annotated or predicted including in *M. musculus*. A BLAT search of the long human BRD8 isoform against the mouse genome revealed homologies between the C-terminal end of the protein sequence and an open reading frame (4933408B1Rik) immediately downstream of the BRD8 gene in the mouse genome (see Fig.10).

Multiple alignment of BRD8 sequences from several species show

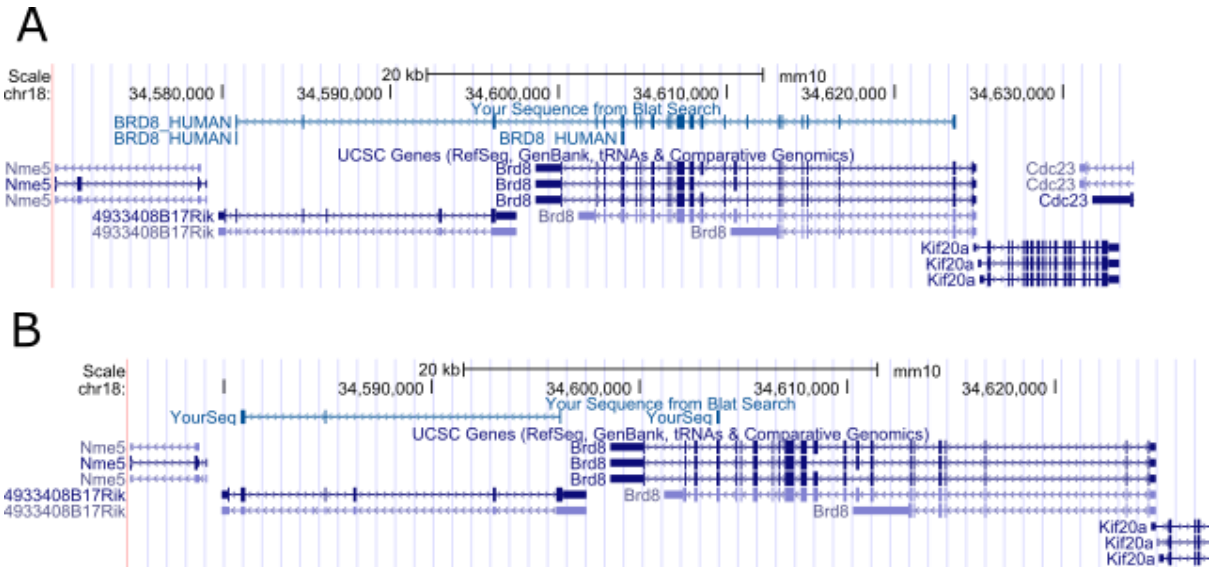

Figure 10: BLAT of long human BRD8 isoform against mouse genome Dec. 2011 (GRCm38/mm10).
A. The whole protein sequence of the long human BRD8 isoform mapped to the mouse genome.
B. The C-terminal of the long human BRD8 isoform, which is non-overlapping with the short human BRD8 isoform, mapped to the mouse genome.

conservation of the first bromodomain (see Fig.11). As for the murine BRD8, the long isoform containing two bromodomains has not been confirmed in a number of the species. Therefore, the second bromodomain does not appear to be conserved. However, a multiple alignment of confirmed long isoform BRD8 proteins reveals conservation of this as well (see Fig.S5).

Furthermore, a segment in the N-terminal end of the protein was also found to have high conservation (see Fig.11). Subsequent BLAST searches and domain prediction for this region came up with no additional information that could elude to the function of this region.

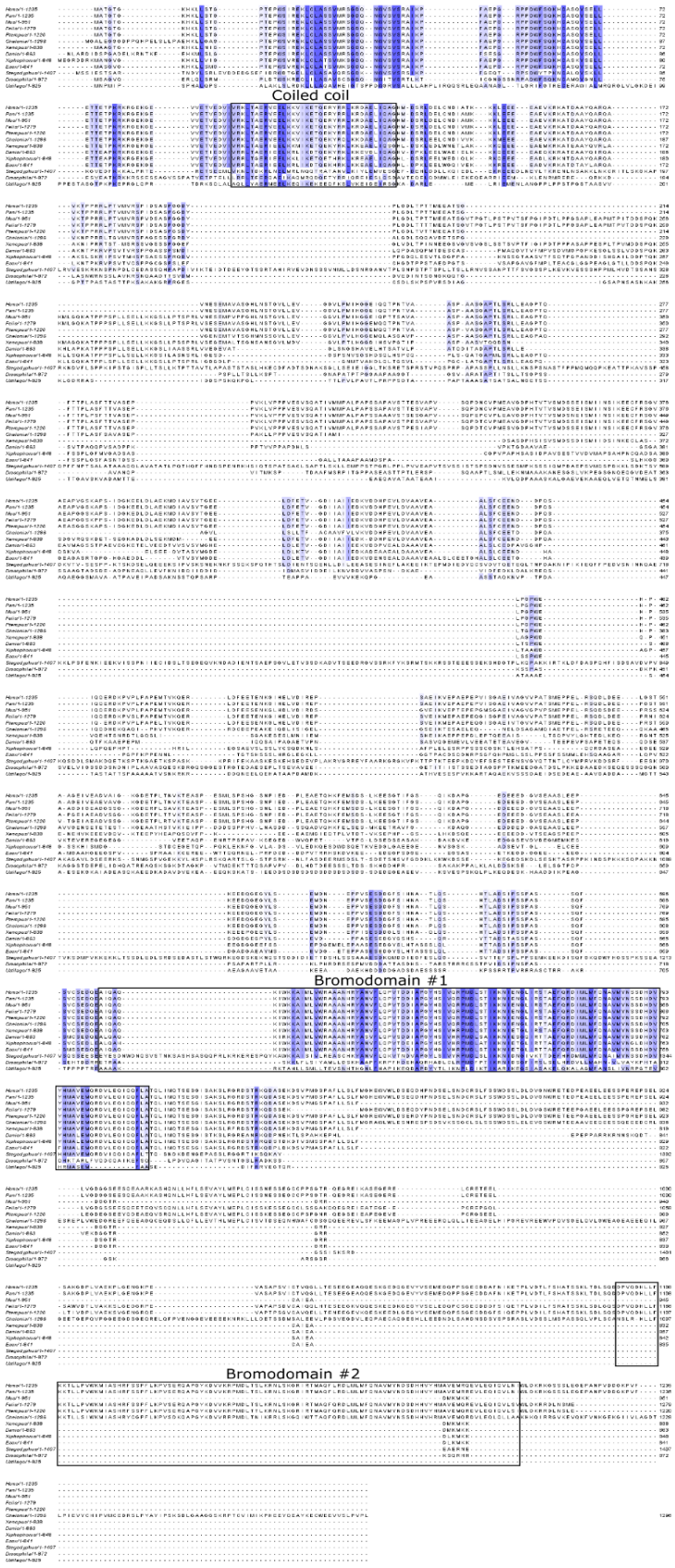

Figure 11: Multiple alignment of BRD8 from Homo sapiens, Pan troglodytes, Mus musculus, Felis catus, Pteropus vampyrus, Chelonia mydas, Xenopus (Silurana) tropicalis, Danio rerio, Xiphophorus maculatus, Esox lucius, Stegodyphus mimosarum, Drosophila melanogaster and Ustilago maydis 521.

The long human isoform of BRD8 has been reported to share some homology with the yeast protein Bdf1 [98]. See supplemental Figure S4 for a pairwise alignment between human BRD8 and *S. cerevisiae* Bdf1. A substantial part of the homology is found in the bromodomains. The human proteins, Brd2 and Brd4, which also have two bromodomains, also share homology with Bdf1.

7.2.3 Both human and mouse BRD8 interact with subunits of the NuA4 HAT complex

The STRING networks for human and murine BRD8 show experimental and text mining evidence for interactions between BRD8 and several components of the NuA4 HAT complex (see Fig.12). This is accordance with the literature (see section 5.2.3.3).

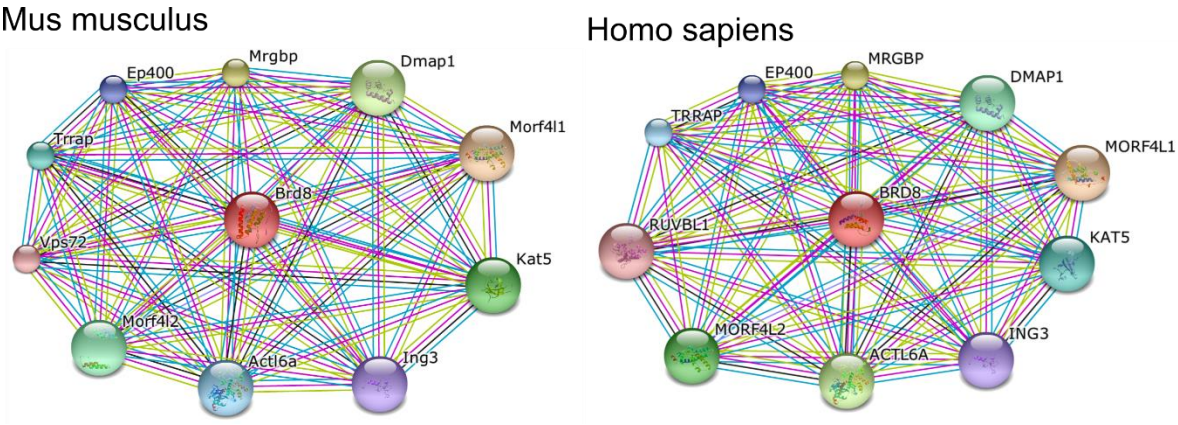

Figure 12: STRING networks for human and murine BRD8.
The 10 interactors with the highest confidence score are shown. The different coloured lines indicate different lines of evidence; black = coexpression, green = textmining, blue = databases and pink = experiments.

The experimental evidence includes co-immunoprecipitation, affinity capture MS-assay, affinity chromatography and affinity capture WB-assay[96,97,120–125]. The interactions shown have a similar STRING confidence score for binding of 0.900. The score is based on experimental data, co-mention in pubmed abstracts and association in curated databases; in this case mainly Gene Ontology.

7.2.4 BRD8 is highly expressed in immune cells

RNA seq data from the BioGPS database reveal that BRD8 is more highly expressed in immune cells compared to other tissue types (see Fig.13). Immgen microarray data focusing exclusively on murine immune cells further show that the highest expression of BRD8 can be found in preB-FrC bone marrow cells and germinal centre cells (see Fig.14). VDJ recombination occurs in the former [17], while SHM and CSR take place in the latter [126].

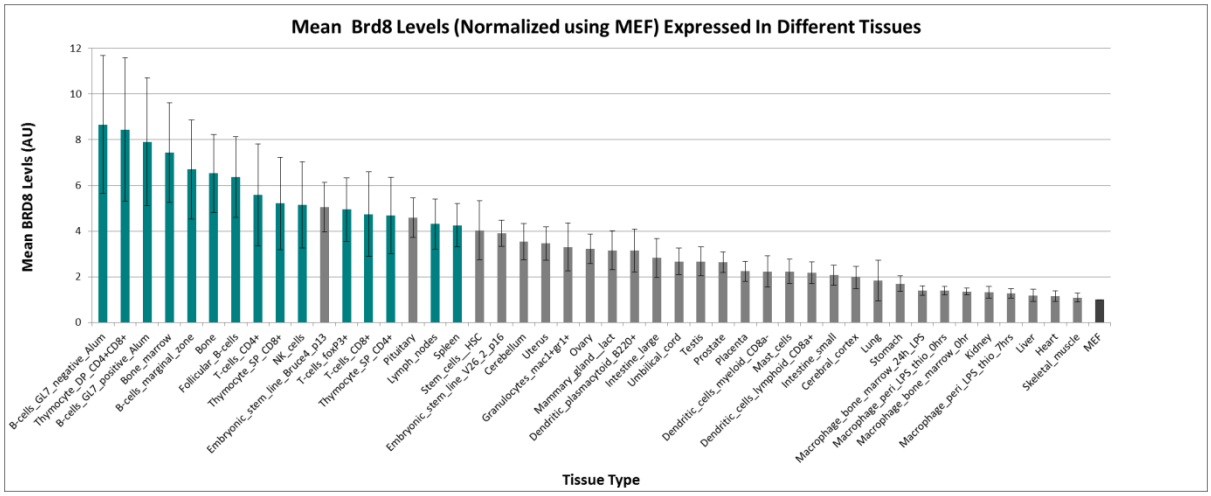

Figure 13: BioGPS RNA seq data expression data.
BRD8 expression level in different tissue types calculated from BioGPS RNA seq data. Immune cells, immune progenitor cells and tissue important for immune cell production or antibody diversification have been marked in green.

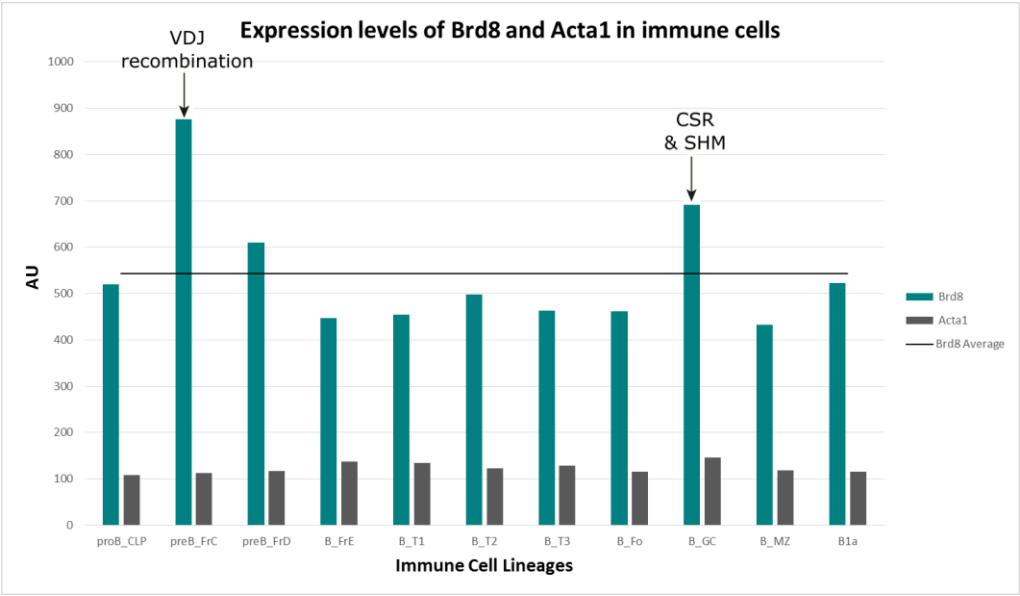

Figure 14: ImmGen microarray expression data.
BRD8 (green) and Acta1 (dark grey) expression levels in immune cells based on microarray data. The average BRD8 expression level is marked by a black line.

7.2.5 BRD8 is localised in the nucleus

Data from Protein Atlas shows that BRD8 is localised in the nucleus and mitochondria of U-2OS cells[111]. Endogenous BRD8 was visualised using the antibody HPA001841 (see Fig.15). BRD8 was not found in the nucleoli[111].

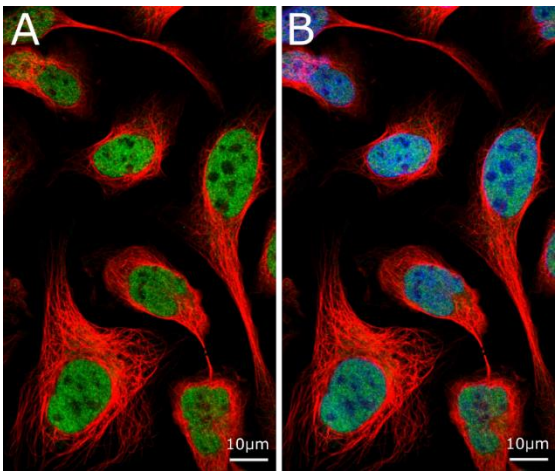

Figure 15: From Protein Atlas. Visualisation of BRD8 subcellular localization in U-2OS cells using immunofluorescence.
BRD8 is visualised in green with the antibody HPA001841 (A and B), while the microtubules are visualised in red (A and B) and the nucleus is visualised in blue (B)[111].

In accordance with this, my own immunofluorescence experiments with camptothecin treated 293A cells showed an accumulation of endogenous BRD8 in the nucleus (see Fig.17A-D). However, background signal in the cytosol was visible. Cells were also stained for the DNA damage mark γH2AX. BRD8 was present in the nucleus of both damaged and undamaged cells.

7.2.6 Verification of BRD8 deletion in transfected CH12 and 3T3 cells

Potential BRD8 deletions caused by the CRISPR-Cas9 system were verified through two approaches; the CJE assay and immunofluorescence. For the 3T3 cells, the gel of the CJE assay did not reveal any additional bands below the original PCR band (see Fig.16). The same was seen for DNA from CH12 cells (data not shown). Deletions in either cell line could not be confirmed based on the CJE assay.

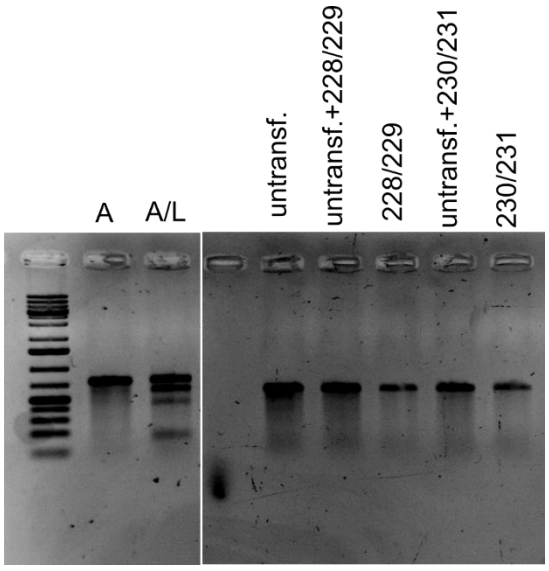

Figure 16: CJE of 3T3 cells transfected with BRD8 gRNAs.*
Control DNA 'A' and 'L' confirm the activity of the CJE assay.
DNA from untransfected 3T3 cells were used for control. Two assays were done for each of the two gRNAs; one alone and one mixed with control DNA at a 1:1 ratio.
**intervening lanes have been cropped out.*

In 3T3 cells, BRD8 knockout was also examined using immunofluorescence. Cells were treated with camptothecin to induce DNA damage and stained with BRD8 and γH2AX antibodies (see Fig.17E-H). As with

the 293A cells, mainly the nucleus of the cells were stained for BRD8. Both damaged and undamaged cells contained BRD8. No nucleus was found without BRD8.

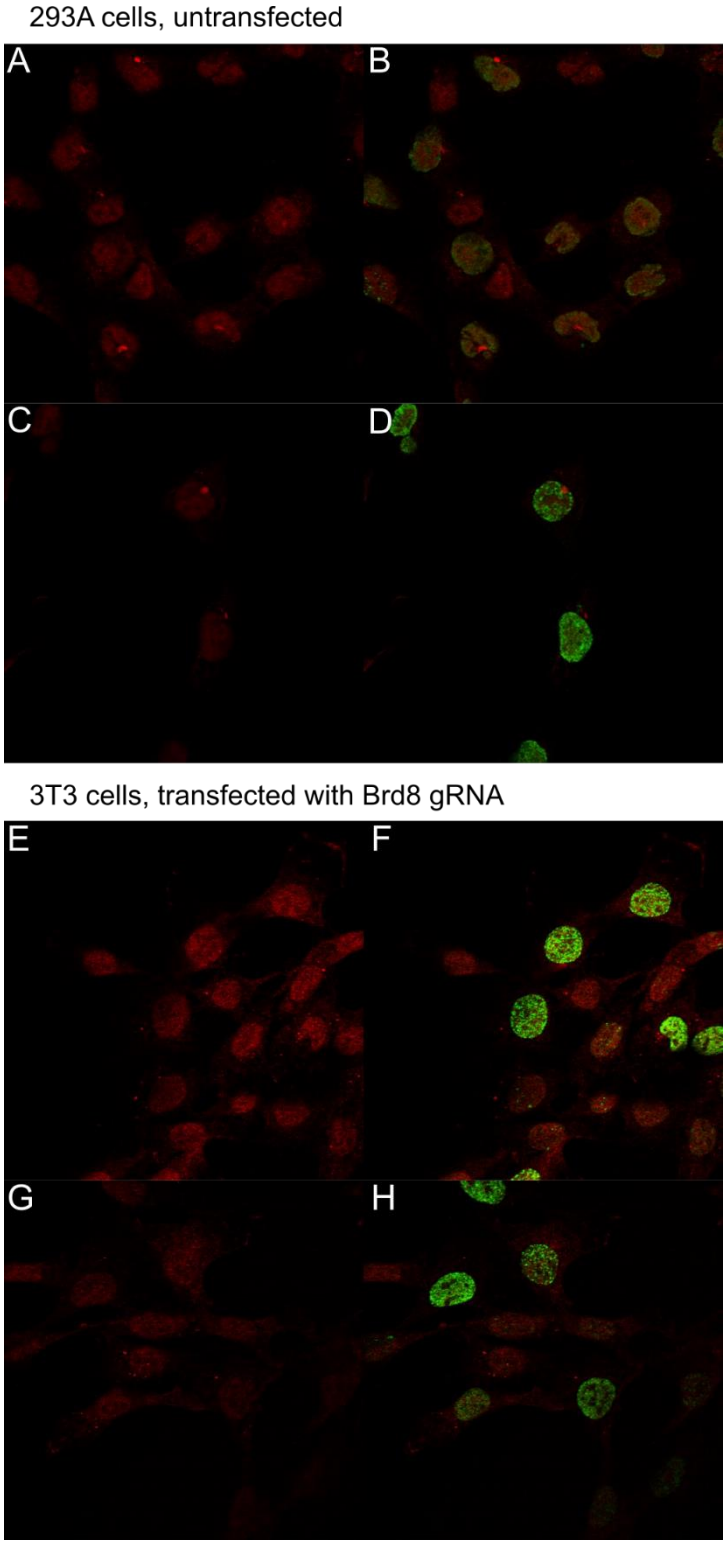

Figure 17: Immunofluorescence of 293A and 3T3 cells stained with BRD8 antibody (red) and γH2AX antibody (green).
All cells were treated with camptothecin for 45 min prior to fixing and staining.

8 Discussion

8.1 Developing a modified CRISPR-Cas9 system

8.1.1 Adar-m2Cas9 expressing plasmid successfully constructed

PCR and sequencing confirmed the direction and sequence of the Adar1 deaminase insert after ligation into the m2Cas plasmid. By using the m2Cas9 plasmid as an expression vector, the expressed fusion protein was the Adar1-dCas9 without the catalytic activity of the wild-type Cas9.

Based on other studies using the Adar1 deaminase domain to target RNA in *trans*, the deaminase activity of Adar1 should be preserved [59]. Thus, the construct is expected to be active and capable of deaminating adenosines of suitable substrates.

8.1.2 Four gRNAs targeting the eGFP gene in 293T/GFP-puro cells provided rapid estimates of mutations efficiencies of Cas9 expressing vectors

In collaboration with Laurence Higgins, I established a suitable model for rapid testing of various Cas9 variants. The model was based on a 293T/GFP-puro cell line continuously expressing the eGFP gene, and four gRNAs targeting this gene. Flow cytometry facilitated rapid and easy assessment of the mutation efficiency of Cas9 variants. Cells transfected with wild-type Cas9 were measured to consist of two subpopulations; one negative and one positive for GFP.

The CJE assay allowed further assessment of the Cas9 variants at the DNA level. The control DNA, with one mismatch in the heteroduplex, validated the activity of the CJE. For the cells transfected with wild-type Cas9 at least one additional band is apparent in the gel. Four gRNAs were transfected into each cell population. For this reason, the heteroduplexes formed before digest with the CJE are expected to contain a variation of mismatches. Additional bands could simply be too small or too few to be detected by the CJE.

Further optimization of the CJE assay will be needed in the future to help detect lower deletion efficiency in transfected cell populations. The use of SYBR Gold for more sensitive detection of DNA may be enough to allow this. The 293T/GFP-puro provides a suitable system for testing this, as deletion percentages can be assessed with flow cytometry prior to analysis with the CJE assay.

8.1.3 The Adar-m2Cas9 construct did not induce measureable mutations in the eGFP gene

The small shift in fluorescence seen for cells transfected with m2Cas9 can be contributed to interference with the transcriptional machinery. This is therefore considered as background and used as a control. All measurements from Adar1 transfected cells should be compared to this control. Potential mutations caused by the Adar1-m2Cas9 construct were not detected by flow cytometry or the CJE assay. There was no significant difference between the fluorescence measured for cells transfected with Adar1-m2Cas9 and cells transfected with m2Cas9 alone.

It is possible that Adar1-m2Cas9 has deaminated adenosines in a small subset of cells making it difficult to measure. Even in the case of high efficiency, it must be considered that the Adar1-m2Cas9 may not deaminate the Tyr-66 codon adenosine. As previously established, the Adar1-m2Cas9 cannot cause stop mutations. Therefore, deamination of adenosines outside the chromophore may not alter the eGFP fluorescence. This would not be detectable by our flow cytometry assay. An optimisation of the CJE assay may reveal more details in the future. Final confirmation of either of these two scenarios would require deep sequencing.

8.1.4 Increased transfection efficiencies could be the key

The deletion efficiency of eGFP in cells transfected with wild-type Cas9 varied depending on the DNA-to-cell ratio. Furthermore, recent transfections using Lipofectamine™ 3000 reagent (Fisher Scientific) increased deletion efficiency as well to more than 90%. This indicates that transfection efficiency is highly important for the resultant deletion efficiency. It seems likely that increased transfection efficiencies will also influence the efficiency of the Adar1-m2Cas9. Future transfections with this and other modified Cas9 systems should be done with Lipofectamine™ 3000.

8.1.5 Mismatches could be necessary for Adar1 activity

Mismatches at or nearby the targeted adenosine increases Adar1 efficiency when targeting dsRNA. It is possible that mismatches would also increase the efficiency of

Adar1 acting on DNA. If Adar1 does indeed target DNA in somatic hypermutation, the initial mismatches caused by AID may serve to recruit it. In the attempt to mimic this system with Adar1-m2Cas9, perhaps a sequential approach for targeting could be the way forward. Whereby the initial mismatches generated by another mutator protein may increase the mutation efficiency of Adar1-m2Cas9.

8.1.6 Adar1-dCas9 as a molecular tool

Optimisation of the Adar-dCas9 system would allow its use in genome engineering. There are three different potential uses for this system.

First, it could assist in the production and diversification of antibodies *ex vivo*. Currently studies of B cell development and antibody diversification, as well as the production of antibodies for research and clinical use, largely rely on whole animal models [127–130]. Recent advances have been made by engineering human immune organoids capable of driving the germinal centre reactions [131]. However, as of yet the processes underlying antibody maturation are not fully understood and the processes cannot be emulated in cell culture. The Adar1-dCas9 could be utilised to circumvent these poorly understood processes and artificially create somatic hypermutation in the variable region of the Ig gene.

Second, it would allow for base-specific mutations *in situ*. The current CRISPR-Cas9 system can only create DSBs. As such it is limited by the NHEJ repair system of the targeted organism. Creating specific point mutations requires a template and induction of the homologous repair system of the targeted organism [132]. The Adar1-dCas9 could potentially simplify this, by targeting specific adenosines guided by the gRNA-dCas9 complex.

Third, the Adar1-dCas9 could be used to change the epigenetic landscape of specific DNA regions. Recently, it was shown that parts of the *C. elegans* genome is methylated on N6-Adenine [133–135]. N6-adenine methylation marks are also found on mRNAs in other eukaryotes [136]. A-to-I deamination by Adar1 would result in a G:T mismatch. If correctly repaired by MMR, a non-methylated adenine would be inserted in the place of the inosine – thus removing the original methylation mark. As such, the Adar1-dCas tool may prove useful for elucidating the role of these methylation marks.

8.2 Assessing the role of BRD8 in antibody diversification and DNA damage repair

8.2.1 BRD8 is highly conserved in vertebrates

BRD8 is highly conserved in vertebrates, and homologs have been established as distantly as *U.maydis*. The fission yeast protein Bdf1 has been implicated as a possible BRD8 homolog. Bdf1 has been implicated in DNA damage repair [137]. However, the limited homology between the two can mainly be found in the bromodomains. Furthermore, other human bromodomain-containing proteins share a similar homology with Bdf1. Therefore, it seems unlikely that Bdf1 is a true homolog of BRD8.

The high conservation of BRD8 is indeed agreement with an essential role, potentially in DNA damage repair and/or transcriptional regulation as implied by its interaction partners (see section 8.2.2). A role in antibody diversification is certainly possible, but cannot be inferred based on sequence conservation alone.

As expected the two bromodomains of BRD8 are highly conserved. A further conserved region was found at the N-terminal end of the protein. This region is of great interest as it may help uncover specific roles for BRD8. It could be an important interaction site or perhaps even a novel domain akin to the bromodomain. Currently, no tested domain predictor recognises it. The fission yeast protein Bdf1 does not contain a region homologous to this conserved region. The discoveries of true BRD8 homologs in yeast may help unravel the role of this region.

8.2.2 Multiple interaction partners could exist for BRD8

The BRD8 sequence contains several predicted low-complexity regions and a coiled-coil region. These have the potential to mediate interactions with other proteins. There is a potential for BRD8 to have several interaction partners. Interaction data summarised in a STRING network supports this notion; indicating that BRD8 interacts with multiple proteins including other subunits of the NuA4 HAT complex. The literature has further implicated BRD8 as a nuclear receptor coactivator [95].

Further studies to reveal BRD8 interactions partners would also shed light on the role of BRD8. Of the established interactions BRD8 interaction partners, Kat5 is

the most interesting. Kat5 has itself been implicated directly in DNA damage repair [85], and indirectly in antibody diversification through the role of the DNA damage pathway in CSR (see section 5.2.3). Furthermore, Kat5 is an acetyltransferase. BRD8 could possibly work downstream of Kat5 mediated acetylation. However, a direct interaction between the two could suggest that BRD8 works in conjunction with Kat5. Perhaps it anchors it to acetylated regions in order to amplify the acetylation signal. The options are many. Unravelling the interaction pattern between the two is certainly important for understanding the role of BRD8 in DNA damage repair and antibody diversification.

8.2.3 Expression data and shRNA screen results support a role in antibody diversification for BRD8

BRD8 was a hit in a genome-wide shRNA screen for proteins involved in CSR. This firmly implicates BRD8 in CSR, and thus antibody diversification. BRD8 expression data supports this. BRD8 is highly expressed in immune cells, or immune system related cells, compared to other cell types. It is further elevated in two cell types; germinal centre cells and pre-B cell bone marrow cells. These are the cell types where SHM, CSR, and VDJ recombination occur, respectively. In short, these are the cells where antibody diversification takes place. The relatively high expression levels in these cell types suggests BRD8 may be involved antibody diversification.

It seems higher Ig DNA damage correlates with higher BRD8 expression. Many proteins are involved in both general DNA damage repair and antibody diversification. Indeed, DSB repair is an essential step in CSR. Determining the role of BRD8 in either process is likely to reveal its role in the other. It is of course possible that BRD8 is involved in several steps of either process. Indeed, the histone mark that recruits Kat5 to DSBs is found upstream of the AID-initiated DSBs in CSR.

8.2.4 BRD8 localises to the nucleus but does not form foci in response to DNA damage

BRD8 localises to the nucleus (Fig. 15). Brd8 staining of 3T3 and 293A cells showed BRD8 concentrating in the nucleus (Fig. 17). Some background was seen in the cytosol, so further optimisation is necessary. The background may be due to

nonspecific binding. Lowering the antibody concentration or increasing the concentration of the mild detergent, TritonX, could eliminate this. Localisation of BRD8 in the nucleus is in agreement with its possible roles in DNA damage repair and antibody diversification.

BRD8 was present in both cells with and without γH2AX foci. Also, BRD8 did not form foci in response to DNA damage. This is in agreement with the literature. In fact, it has been shown, that BRD8 does not re-localise after DNA damage by laser microirradiation [93]. It is important to note that this does not refute a role for BRD8 in DNA damage repair. Several DNA damage response genes have been shown to have essential roles in relaying DNA damage signalling without the necessity to localise to DNA damage foci, e.g. CHK1 and CHK2 protein kinases [138]. BRD8, as part of the NuA4 HAT complex, is also expected to be involved in transcriptional activation. For this reason, it is likely that BRD8 is always needed in the cell at a relatively high level. Small changes in protein levels and localisation within the nucleus may not be visible by confocal.

8.2.5 CRISPR-Cas9 mediated knockout of BRD8

The efficiency of the CRISPR-Cas9 system was confirmed in 293T/GFP-puro cells with eGFP targeting gRNAs. Unfortunately, BRD8 deletions could not be confirmed in 3T3 or CH12 cells. Whether this is due to lack of deletion or shortcomings in the assay methods is not known. Further optimisation of both the CJE assay and the BRD8 antibody staining may reveal more.

Another option must also be considered. It is possible that BRD8 knockout is lethal for the cells. Indeed, siRNA knockdown of BRD8 has been shown to affect different cells differently, with an increase in cell death for some cell lines [94].

9 Future directions

9.1 Developing a modified CRISPR-Cas9 system

9.1.1 Mismatches and transcription may be required for Adar1 activity

The Adar1-Cas9 transfections into 293T/GFP-puro will be optimised. The eGFP gene provides an excellent target for this with easily measurable results. Optimisation of the CJE assay alongside should increase ease the analysis of the Adar1-Cas9 effect on the eGFP locus.

Current work includes a sequential approach with another Cas9-fusion protein capable of creating mismatches. Preliminary results from these double-transfections are positive. The eGFP signal is shifting more than for either Cas9-fusion alone. Induction of transcription in the targeted region also has an effect. The role of transcription is unknown. The Adar1-m2Cas9 does not contain the Z-DNA binding domain, so the induction of Z-DNA formation should not play a role. Whether Adar1 is exerting its deamination on the transcribed RNA or perhaps the ssDNA is not known. Further studies are needed and deep sequencing of the targeted region should reveal more.

Changes to the Adar1-m2Cas9 construct may also be necessary for optimal efficiency. Fusion of dCas9 to the N-terminus of Adar1 or changes to the linker between the two should be considered.

9.1.2 Elucidate the role of Adar1 in antibody diversification *in vivo*

The imitation of SHM through the CRISPR-Cas9 system is interesting on several levels. When optimised the construct has several potential uses, and it may assist in unravelling the role of Adar1 in SHM *in vivo*. However, for a full understanding of this role, *in vivo* assays in suitable models are necessary. Knockout or inhibition of Adar1 could prove useful in determining the effect of Adar1. Separating its canonical dsRNA activity from its potential DNA activity may prove difficult. Here, the synthetic system may be more suitable. A decrease in GFP measured by GFP without the corresponding mutations being visible in an optimised CJE assay, would indicate Adar1 action on RNA, rather than DNA.

In this thesis, I have attempted to utilise the deamination capability of Adar1 to imitate SHM *ex vivo*. That is not to say that it is not involved in both SHM and CSR *in vivo*. Indeed, the Z-DNA binding domain of Adar1 has been shown to bind G-quadruplexes [139]. This interaction mediates the recruitment of Adar1 to the myc promotor [139]. Recently, G-quadruplexes were shown to mediate the recruitment of AID to the switch regions [61]. In the absence of switch region G-quadruplexes CSR was impaired [61]. Adar1 could potentially be recruited to switch regions in a similar way.

9.2 Assessing the role of BRD8 in antibody diversification and DNA damage repair

9.2.1 Identify potential BRD8 homologs in yeast

Although BRD8 shares homology with yeast Bdf1 [98], the evolutionary relationship between the two is still questionable. Focusing on the N-terminal conserved region, further searches will be carried out in an attempt to find true yeast homologs. This region is of particular interest in terms of BRD8 functions, and homologs in yeast may help uncover this.

9.2.2 Verify BRD8 deletions in mouse cells

Mouse cells are a suitable model for antibody diversification and DNA damage repair because of their close relation with human cells. However, it is important to take into consideration, that only the short isoform of BRD8, containing a single bromodomain, has been annotated in mouse. Therefore, the murine BRD8 will be characterized further. The possibly incorrectly annotated region downstream of (see section 7.2.2), and available RNA sequencing data from mouse cells will aid this characterization.

The BRD8 deletions in 3T3 and CH12 cells will be verified. Optimisation of the CJE assay for better visualisation of fainter bands could potentially reveal deletions that have not been verified as of yet. Further optimisation of the IF assay could do the same. The presence or absence of the BRD8 protein will also be assayed using western blot.

9.2.3 Isolate BRD8 deletion clones for 3T3 and CH12 cells

Upon verification of deletion the cells will be diluted. Single clones will be picked and BRD8 deletion verified with CJE and IF. Sequencing of these clones will then confirm the exact nature of the knockout.

If BRD8 deletion is lethal for the cells, other measures need to be taken. A specific inhibitor for BRD8 would be of great use. However, due to the homology between bromodomain-containing proteins, development of such an inhibitor is difficult. Knock-out of BRD8 followed by transfection with an inducible BRD8-expression system is a more viable option. This would allow for controlled inhibition of BRD8 expression. Alternatively, siRNA knockdown may be an option. Up to 95% knockdown has been achieved [94].

9.2.4 Assess DNA damage response in ΔBRD8 fibroblast cell lines

DNA damage will be induced through camptothecin and irradiation and the DNA damage response will be assessed for wild-type cells and ΔBRD8 cells. The accumulation of γH2AX at DSBs will be assayed using flow cytometry and IF. In this way, I can determine if BRD8 functions upstream or downstream of the γH2AX mark. Based on its interaction with Kat5, it is expected to function alongside Kat5 upstream of γH2AX.

9.2.5 Assess antibody diversification in ΔBRD8 CH12 cells

CH12 cells are capable of switching from IgM to IgA, so they provide an excellent model for studying CSR. The proportion of IgA switched cells within a population will be assayed using fluorescent antibodies and flow cytometry. Comparison of wild-type cells and ΔBRD8 cells will help unravel the role of BRD8 in CSR. Combination assays with Kat5 inhibition in ΔBRD8 cells may reveal more; both about the BRD8-Kat5 interaction and their roles in antibody diversification.

9.2.6 Deletion of BRD8 in human cells

While murine cells are an excellent model for DNA damage repair and antibody diversification, human cells may reveal additional information. gRNAs have been designed for the human BRD8. Using the CRISPR-Cas9 system I will attempt to

knock-out the BRD8 gene in human 293A fibroblast cells. These can be used for modelling DNA damage repair and help elucidate the role of BRD8.

10 Glossary

ADAR	Adenosine deaminase acting on RNA
AID	Activation-induced cytidine deaminase
Bdf1	Bromodomain-containing Factor 1
BRD8/BRD8	Bromodomain containing 8 (mouse/human)
CJE	Celery Juice Extract
crRNA	CRISPR RNA
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSR	Class-Switch Recombination
DSB	Double-strand break
gRNA	guideRNA
HR	Homologous recombination
NHEJ	Non-homologous end joining
PEI	Polyethylenimine
SHM	Somatic Hypermutation
tracrRNA	trans-activating crRNA

11 References

1. Tsuruoka, N. et al. 2013 ADAR1 protein induces adenosine-targeted DNA mutations in senescent Bcl6 gene-deficient cells. *J. Biol. Chem.* **288**, 826–836. (doi:10.1074/jbc.M112.365718)

2. Horvath, P. & Barrangou, R. 2010 CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167–170. (doi:10.1126/science.1179555)

3. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. a & Horvath, P. 2007 CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science (80-.).* **315**, 1709–1712.

4. Cong, L. et al. 2013 Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–23. (doi:10.1126/science.1231143)

5. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. 2013 RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239. (doi:10.1038/nbt.2508)

6. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. 2012 Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E2579–86. (doi:10.1073/pnas.1208507109)

7. Zhang, Y., Heidrich, N., Ampattu, B. J., Gunderson, C. W., Seifert, H. S., Schoen, C., Vogel, J. & Sontheimer, E. J. 2013 Processing-Independent CRISPR RNAs Limit Natural Transformation in Neisseria meningitidis. *Mol. Cell* **50**, 488–503. (doi:10.1016/j.molcel.2013.05.001)

8. Ran, F., Hsu, P., Wright, J. & Agarwala, V. 2013 Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–308. (doi:10.1038/nprot.2013.143)

9. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. 2012 A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.).* **337**, 816–822. (doi:10.1126/science.1225829)

10. Naito, Y., Hino, K., Bono, H. & Ui-Tei, K. 2014 CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* **31**, 1120–1123. (doi:10.1093/bioinformatics/btu743)

11. Guilinger, J. P., Thompson, D. B. & Liu, D. R. 2014 Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* **32**, 577–82. (doi:10.1038/nbt.2909)

12. Deng, W., Shi, X., Tjian, R., Lionnet, T. & Singer, R. H. 2015 CASFISH : CRISPR / Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc. Natl. Acad. Sci.* **112**, 11870–11875. (doi:10.1073/pnas.1515692112)

13. Chevez, A. et al. 2015 Highly-efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**, 326–328. (doi:10.1038/nmeth.3312.Highly-efficient)

14. Lawhorn, I. E. B., Ferreira, J. P. & Wang, C. L. 2014 Evaluation of sgRNA

target sites for CRISPR-mediated repression of TP53. *PLoS One* **9**, e113232. (doi:10.1371/journal.pone.0113232)

15. Parra, D., Takizawa, F. & Sunyer, J. O. 2013 Evolution of B Cell Immunity. *Annu. Rev. Anim. Biosci.* **1**, 65–97. (doi:10.1016/j.micinf.2011.07.011.Innate)

16. Vuong, B. Q. & Chaudhuri, J. 2012 Combinatorial mechanisms regulating AID-dependent DNA deamination: interacting proteins and post-translational modifications. *Semin. Immunol.* **24**, 264–72. (doi:10.1016/j.smim.2012.05.006)

17. Johnson, K., Chaumeil, J. & Skok, J. A. 2010 Epigenetic regulation of V(D)J recombination. *Essays Biochem.* **48**, 221–243. (doi:10.1042/bse0480221)

18. Martincorena, I. & Campbell, P. J. 2015 Somatic mutation in cancer and normal cells. *Science (80-.).* **349**, 1483–1489.

19. Kracker, S. & Durandy, A. 2011 Insights into the B cell specific process of immunoglobulin class switch recombination. *Immunol Lett* **138**, 97–103. (doi:10.1016/j.imlet.2011.02.004)

20. Odegard, V. H. & Schatz, D. G. 2006 Targeting of somatic hypermutation. *Nat. Rev. Immunol.* **6**, 573–583. (doi:10.1038/nri1896)

21. Peled, J. U., Kuang, F. L., Iglesias-Ussel, M. D., Roa, S., Kalis, S. L., Goodman, M. F. & Scharff, M. D. 2008 The biochemistry of somatic hypermutation. *Annu. Rev. Immunol.* **26**, 481–511. (doi:10.1146/annurev.immunol.26.021607.090236)

22. Di Noia, J. M. & Neuberger, M. S. 2007 Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* **76**, 1–22. (doi:10.1146/annurev.biochem.76.061705.090740)

23. Dickerson, S. K., Market, E., Besmer, E. & Papavasiliou, F. N. 2003 AID mediates hypermutation by deaminating single stranded DNA. *J. Exp. Med.* **197**, 1291–1296. (doi:10.1084/jem.20030481)

24. Wiesendanger, M., Kneitz, B., Edelmann, W. & Scharff, M. D. 2000 Somatic hypermutation in MutS homologue (MSH)3-, MSH6-, and MSH3/MSH6-deficient mice reveals a role for the MSH2-MSH6 heterodimer in modulating the base substitution pattern. *J. Exp. Med.* **191**, 579–84. (doi:10.1084/jem.191.3.579)

25. Rada, C., Ehrenstein, M. R., Neuberger, M. S. & Milstein, C. 1998 Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity* **9**, 135–41. (doi:10.1016/S1074-7613(00)80595-6)

26. Frieder, D., Larijani, M., Collins, C., Shulman, M. & Martin, A. 2009 The concerted action of Msh2 and UNG stimulates somatic hypermutation at A . T base pairs. *Mol. Cell. Biol.* **29**, 5148–5157. (doi:10.1128/MCB.00647-09)

27. Saribasak, H., Rajagopal, D., Maul, R. W. & Gearhart, P. J. 2009 Hijacked DNA repair proteins and unchained DNA polymerases. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**, 605–611. (doi:10.1098/rstb.2008.0188)

28. Bransteitter, R., Pham, P., Scharff, M. D. & Goodman, M. F. 2003 Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded

DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 4102–4107. (doi:10.1073/pnas.0730835100)

29. Kano, C. & Wang, J. Y. 2013 High levels of AID cause strand bias of mutations at A versus T in Burkitt's lymphoma cells. *Mol. Immunol.* **54**, 397–402. (doi:10.1016/j.molimm.2013.01.005)

30. Franklin, A., Milburn, P. J., Blanden, R. V & Steele, E. J. 2004 Human DNA polymerase-eta, an A-T mutator in somatic hypermutation of rearranged immunoglobulin genes, is a reverse transcriptase. *Immunol. Cell Biol.* **82**, 219–225. (doi:10.1046/j.0818-9641.2004.01221.x)

31. Masutani, C., Kusumoto, R., Iwai, S. & Hanaoka, F. 2000 Mechanisms of accurate translesion synthesis by human DNA polymerase eta. *EMBO J.* **19**, 3100–3109. (doi:10.1093/emboj/19.12.3100)

32. Johnson, R. E., Washington, M. T., Prakash, S. & Prakash, L. 2000 Fidelity of human DNA polymerase eta. *J. Biol. Chem.* **275**, 7447–7450. (doi:10.1074/jbc.275.11.7447)

33. Yavuz, S., Yavuz, A. S., Kraemer, K. H. & Lipsky, P. E. 2002 The role of polymerase eta in somatic hypermutation determined by analysis of mutations in a patient with xeroderma pigmentosum variant. *J. Immunol.* **169**, 3825–30. (doi:10.4049/jimmunol.169.7.3825)

34. Zeng, X., Winter, D. B., Kasmer, C., Kraemer, K. H., Lehmann, A. R. & Gearhart, P. J. 2001 Human DNA polymerase-eta is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat. Immunol.* **2**, 537–541. (doi:10.1046/j.0818-9641.2004.01221.x)

35. Rogozin, I. B., Pavlov, Y. I., Bebenek, K., Matsuda, T. & Kunkel, T. a 2001 Somatic mutation hotspots correlate with DNA polymerase eta error spectrum. *Nat. Immunol.* **2**, 530–536. (doi:10.1038/88732)

36. Pavlov, Y. I., Rogozin, I. B., Galkin, A. P., Aksenova, A. Y., Hanaoka, F., Rada, C. & Kunkel, T. A. 2002 Correlation of somatic hypermutation specificity and A-T base pair substitution errors by DNA polymerase eta during copying of a mouse immunoglobulin kappa light chain transgene. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 9954–9. (doi:10.1073/pnas.152126799)

37. Delbos, F., De Smet, A., Faili, A., Aoufouchi, S., Weill, J.-C. & Reynaud, C.-A. 2005 Contribution of DNA polymerase eta to immunoglobulin gene hypermutation in the mouse. *J. Exp. Med.* **201**, 1191–1196. (doi:10.1084/jem.20050292)

38. Zhao, Y., Gregory, M. T., Biertümpfel, C., Hua, Y., Hanaoka, F. & Yang, W. 2013 Mechanism of somatic hypermutation at the WA motif by human DNA polymerase η. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 8146–51. (doi:10.1073/pnas.1303126110)

39. Steele, E. J. 2009 Mechanism of somatic hypermutation: Critical analysis of strand biased mutation signatures at A:T and G:C base pairs. *Mol. Immunol.* **46**, 305–320. (doi:10.1016/j.molimm.2008.10.021)

40. Franklin, A. & Blanden, R. V 2006 A/T-targeted somatic hypermutation: critique of the mainstream model. *Trends Biochem. Sci.* **31**, 252–8.

(doi:10.1016/j.tibs.2006.03.008)

41. Franklin, A. & Blanden, R. V 2008 The strand bias paradox of somatic hypermutation at immunoglobulin loci. *Trends Immunol.* **29**, 167–72. (doi:10.1016/j.it.2008.01.008)

42. Mayorov, V. I., Rogozin, I. B., Adkison, L. R. & Gearhart, P. J. 2005 DNA polymerase eta contributes to strand bias of mutations of A versus T in immunoglobulin genes. *J. Immunol.* **174**, 7781–6. (doi:174/12/7781)

43. Zivojnovic, M., Delbos, F., Girelli Zubani, G., Julé, A., Alcais, A., Weill, J.-C., Reynaud, C.-A. & Storck, S. 2014 Somatic hypermutation at A/T-rich oligonucleotide substrates shows different strand polarities in Ung-deficient or -proficient backgrounds. *Mol. Cell. Biol.* **34**, 2176–87. (doi:10.1128/MCB.01452-13)

44. Wong, S. K., Sato, S. & Lazinski, D. W. 2001 Substrate recognition by ADAR1 and ADAR2. *RNA* **7**, 846–58. (doi:10.1017/S135583820101007X)

45. Steele, E. J., Lindley, R. A., Wen, J. & Weiller, G. F. 2006 Computational analyses show A-to-G mutations correlate with nascent mRNA hairpins at somatic hypermutation hotspots. *DNA Repair* **5**, 1346–1363. (doi:10.1016/j.dnarep.2006.06.002)

46. Patterson, J. B. & Samuel, C. E. 1995 Expression and Regulation by Interferon of a Double-Stranded- RNA-Specific Adenosine Deaminase from Human Cells : Evidence for Two Forms of the Deaminase. **15**, 5376–5388.

47. Patterson, J. B., Thomis, D. C., Hans, S. L. & Samuel, C. E. 1995 Mechanism of interferon action: double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons. *Virology* **210**, 508–11. (doi:10.1006/viro.1995.1370)

48. Kawakubo, K. & Samuel, C. E. 2000 Human RNA-specific adenosine deaminase (ADAR1) gene specifies transcripts that initiate from a constitutively active alternative promoter. *Gene* **258**, 165–172. (doi:10.1016/S0378-1119(00)00368-1)

49. Barraud, P. 2012 ADAR Proteins: Double-stranded RNA and Z-DNA Binding Domains. *Curr. Top. Microbiol. Immunol.* **351**, 139–157. (doi:10.1007/82)

50. Choi, J. & Majima, T. 2011 Conformational changes of non-B DNA. *Chem. Soc. Rev.* **40**, 5893. (doi:10.1039/c1cs15153c)

51. Liu, L. F. & Wang, J. C. 1987 Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 7024–7027. (doi:10.1073/pnas.84.20.7024)

52. Peck, L. J., Nordheim, A., Rich, A. & Wang, J. C. 1982 Flipping of cloned d(pCpG)n.d(pCpG)n DNA sequences from right- to left-handed helical structure by salt, Co(III), or negative supercoiling. *Proc Natl Acad Sci U S A* **79**, 4560–4564. (doi:10.1073/pnas.79.15.4560)

53. Wong, C. W. & Privalsky, M. L. 1998 Components of the SMRT corepressor complex exhibit distinctive interactions with the POZ domain oncoproteins PLZF, PLZF-RARalpha, and BCL-6. *J. Biol. Chem.* **273**, 27695–702.

(doi:10.1074/jbc.273.42.27695)

54. Shvarts, A., Brummelkamp, T. R., Scheeren, F., Koh, E., Daley, G. Q., Spits, H. & Bernards, R. 2002 A senescence rescue screen identifies BCL6 as an inhibitor of anti-proliferative p19ARF-p53 signaling. *Genes Dev.* **16**, 681–686. (doi:10.1101/gad.929302.GENES)

55. Dent, A. L., Shaffer, A. L., Yu, X., Allman, D. & Staudt, L. M. 1997 Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science* **276**, 589–592. (doi:10.1126/science.276.5312.589)

56. Ye, B. H. et al. 1997 The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nat. Genet.* **16**, 161–70. (doi:10.1038/ng0697-161)

57. Fukuda, T. et al. 1997 Disruption of the Bcl6 gene results in an impaired germinal center formation. *J. Exp. Med.* **186**, 439–448. (doi:10.1084/jem.186.3.439)

58. Herbert, A. & Rich, A. 2001 The role of binding domains for dsRNA and Z-DNA in the in vivo editing of minimal substrates by ADAR1. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 12132–12137. (doi:10.1073/pnas.211419898)

59. Schneider, M. F., Wettengel, J., Hoffmann, P. C. & Stafforst, T. 2014 Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. *Nucleic Acids Res.* **42**. (doi:10.1093/nar/gku272)

60. Kenter, A. L. 2012 AID targeting is dependent on RNA polymerase II pausing. *Semin. Immunol.* **24**, 281–6. (doi:10.1016/j.smim.2012.06.001)

61. Zheng, S., Vuong, B. Q., Vaidyanathan, B., Lin, J.-Y., Huang, F.-T. & Chaudhuri, J. 2015 Non-coding RNA Generated following Lariat Debranching Mediates Targeting of AID to DNA. *Cell* **161**, 762–773.

62. Aida, M. & Honjo, T. 2013 FACT and H3.3: New markers for the somatic hypermutation. *Cell Cycle* **12**, 2923–2924. (doi:10.4161/cc.26178)

63. Stanlie, A., Aida, M., Muramatsu, M., Honjo, T. & Begum, N. a 2010 Histone3 lysine4 trimethylation regulated by the facilitates chromatin transcription complex is critical for DNA cleavage in class switch recombination. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 22190–22195. (doi:10.1073/pnas.1016923108)

64. Begum, N. A., Stanlie, A., Nakata, M., Akiyama, H. & Honjo, T. 2012 The histone chaperone Spt6 is required for activation-induced cytidine deaminase target determination through H3K4me3 regulation. *J. Biol. Chem.* **287**, 32415–29. (doi:10.1074/jbc.M112.351569)

65. Begum, N. A., Stanlie, A., Nakata, M., Akiyama, H. & Honjo, T. 2012 The histone chaperone Spt6 is required for activation-induced cytidine deaminase target determination through H3K4me3 regulation. *J. Biol. Chem.* **287**, 32415–32429. (doi:10.1074/jbc.M112.351569)

66. Okazaki, I. et al. 2011 Histone chaperone Spt6 is required for class switch recombination but not somatic hypermutation. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 7920–7925. (doi:10.1073/pnas.1104423108)

67. Kuang, F. L., Luo, Z. & Scharff, M. D. 2009 H3 trimethyl K9 and H3 acetyl K9

chromatin modifications are associated with class switch recombination. *Pnas*

68. Sun, Y., Jiang, X., Xu, Y., Ayrapetov, M. K., Moreau, L. a, Whetstine, J. R. & Price, B. D. 2009 Histone H3 methylation links DNA damage detection to activation of the tumour suppressor Tip60. *Nat. Cell Biol.* **11**, 1376–1382. (doi:10.1038/ncb1982)

69. Gospodinov, A. & Herceg, Z. 2013 Chromatin structure in double strand break repair. *DNA Repair (Amst).* **12**, 800–810. (doi:10.1016/j.dnarep.2013.07.006)

70. Li, G., White, C. a., Lam, T., Pone, E. J., Tran, D. C., Hayama, K. L., Zan, H., Xu, Z. & Casali, P. 2013 Combinatorial H3K9acS10ph histone modification in igh locus s regions targets 14-3-3 adaptors and aid to specify antibody class-switch DNA recombination. *Cell Rep.* **5**, 702–714. (doi:10.1016/j.celrep.2013.09.031)

71. Winter, S., Fischle, W. & Seiser, C. 2008 Modulation of 14-3-3 interaction with phosphorylated histone H3 by combinatorial modification patterns. *Cell Cycle* **7**, 1336–1342. (doi:10.4161/cc.7.10.5946)

72. Walter, W., Clynes, D., Tang, Y., Marmorstein, R., Mellor, J. & Berger, S. L. 2008 14-3-3 interaction with histone H3 involves a dual modification pattern of phosphoacetylation. *Mol. Cell. Biol.* **28**, 2840–2849. (doi:10.1128/MCB.01457-07)

73. Xu, Z. et al. 2010 14-3-3 adaptor proteins recruit AID to 5’-AGCT-3'-rich switch regions for class switch recombination. *Nat. Struct. Mol. Biol.* **17**, 1124–35. (doi:10.1038/nsmb.1884)

74. Noon, A. T. & Goodarzi, A. a. 2011 53BP1-mediated DNA double strand break repair: Insert bad pun here. *DNA Repair (Amst).* **10**, 1071–1076. (doi:10.1016/j.dnarep.2011.07.012)

75. Nakamura, K., Sakai, W., Kawamoto, T., Bree, R. T., Lowndes, N. F., Takeda, S. & Taniguchi, Y. 2006 Genetic dissection of vertebrate 53BP1: A major role in non-homologous end joining of DNA double strand breaks. *DNA Repair (Amst).* **5**, 741–749. (doi:10.1016/j.dnarep.2006.03.008)

76. Dimitrova, N., Chen, Y.-C. M., Spector, D. L. & de Lange, T. 2008 53BP1 promotes non-homologous end joining of telomeres by increasing chromatin mobility. *Nature* **456**, 524–8. (doi:10.1038/nature07433)

77. Ward, I. M. 2004 53BP1 is required for class switch recombination. *J. Cell Biol.* **165**, 459–464. (doi:10.1083/jcb.200403021)

78. Manis, J. P., Morales, J. C., Xia, Z., Kutok, J. L., Alt, F. W. & Carpenter, P. B. 2004 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. *Nat. Immunol.* **5**, 481–7. (doi:10.1038/ni1067)

79. Bothmer, A., Robbiani, D. F., Feldhahn, N., Gazumyan, A., Nussenzweig, A. & Nussenzweig, M. C. 2010 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. *J. Exp. Med.* **207**, 855–65. (doi:10.1084/jem.20100244)

80. Zimmermann, M. & de Lange, T. 2014 53BP1: pro choice in DNA repair.

Trends Cell Biol. **24**, 108–117. (doi:10.1016/j.tcb.2013.09.003)

81. Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J. & Mer, G. 2006 Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair. *Cell* **127**, 1361–1373. (doi:10.1016/j.cell.2006.10.043)

82. Fradet-Turcotte, A. et al. 2013 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature* **499**, 50–56. (doi:10.1038/nature12318)

83. Pei, H., Wu, X., Liu, T., Yu, K., Jelinek, D. F. & Lou, Z. 2013 The histone methyltransferase MMSET regulates class switch recombination. *J. Immunol.* **190**, 756–63. (doi:10.4049/jimmunol.1201811)

84. Hong, S., Dutta, A. & Laimins, L. A. 2015 The Acetyltransferase Tip60 Is a Critical Regulator of the Differentiation-Dependent Amplification of Human Papillomaviruses. *J. Virol.* **89**, 4668–4675. (doi:10.1128/JVI.03455-14)

85. Kleiner, R. E., Verma, P., Molloy, K. R., Chait, B. T. & Kapoor, T. M. 2015 Chemical proteomics reveals a γH2AX-53BP1 interaction in the DNA damage response. *Nat. Chem. Biol.* **11**, 807–814.

86. Dey, A., Chitsaz, F., Abbasi, A., Misteli, T. & Ozato, K. 2003 The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 8758–8763. (doi:10.1073/pnas.1433065100)

87. Stanlie, A., Yousif, A. S., Akiyama, H., Honjo, T. & Begum, N. a. 2014 Chromatin reader Brd4 functions in Ig class switching as a repair complex adaptor of nonhomologous end-joining. *Mol. Cell* **55**, 97–110. (doi:10.1016/j.molcel.2014.05.018)

88. Ramachandran, S. et al. 2010 The RNF8/RNF168 ubiquitin ligase cascade facilitates class switch recombination. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 809–814. (doi:10.1073/pnas.0913790107)

89. Mattiroli, F., Vissers, J. H. a, Van Dijk, W. J., Ikpa, P., Citterio, E., Vermeulen, W., Marteijn, J. a. & Sixma, T. K. 2012 RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. *Cell* **150**, 1182–1195. (doi:10.1016/j.cell.2012.08.005)

90. Gatti, M., Pinato, S., Maspero, E., Soffientini, P., Polo, S. & Penengo, L. 2012 A novel ubiquitin mark at the N-terminal tail of histone H2As targeted by RNF168 ubiquitin ligase. *Cell Cycle* **11**, 2538–2544. (doi:10.4161/cc.20919)

91. Kocyłowski, M. K., Rey, A. J., Stewart, G. S. & Halazonetis, T. D. 2015 Ubiquitin-H2AX fusions render 53BP1 recruitment to DNA damage sites independent of RNF8 or RNF168. *Cell Cycle* , 37–41. (doi:10.1080/15384101.2015.1010918)

92. Sanchez, R. & Zhou, M.-M. 2009 The role of human bromodomains in chromatin biology and gene transcription. *Curr. Opin. Drug Discov. Devel.* **12**, 659–665. (doi:10.1016/j.bbi.2008.05.010)

93. Gong, F. et al. 2015 Screen identifies bromodomain protein ZMYND8 in chromatin recognition of transcription-associated DNA damage that promotes

homologous recombination. *Genes Dev.* **29**, 197–211. (doi:10.1101/gad.252189.114.DNA)

94. Yamada, H. Y. & Rao, C. V. 2009 BRD8 is a potential chemosensitizing target for spindle poisons in colorectal cancer therapy. *Int. J. Oncol.* **35**, 1101–1109. (doi:10.3892/ijo)

95. Monden, T., Wondisford, F. E. & Hollenberg, A. N. 1997 Isolation and characterization of a novel ligand-dependent thyroid hormone receptor-coactivating protein. *J. Biol. Chem.* **272**, 29834–29841. (doi:10.1074/jbc.272.47.29834)

96. Cai, Y., Jin, J., Tomomori-Sato, C., Sato, S., Sorokina, I., Parmely, T. J., Conaway, R. C. & Conaway, J. W. 2003 Identification of New Subunits of the Multiprotein Mammalian TRRAP/TIP60-containing Histone Acetyltransferase Complex. *J. Biol. Chem.* **278**, 42733–42736. (doi:10.1074/jbc.C300389200)

97. Doyon, Y., Selleck, W., Lane, W. S., Tan, S., Côté, J. & Co, J. 2004 Structural and Functional Conservation of the NuA4 Histone Acetyltransferase Complex from Yeast to Humans Structural and Functional Conservation of the NuA4 Histone Acetyltransferase Complex from Yeast to Humans. *Mol. Cell. Biol.* **24**, 1884–96. (doi:10.1128/MCB.24.5.1884)

98. Doyon, Y. & Côté, J. 2004 The highly conserved and multifunctional NuA4 HAT complex. *Curr. Opin. Genet. Dev.* **14**, 147–154. (doi:10.1016/j.gde.2004.02.009)

99. Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L. & Church, G. M. 2013 Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31**, 833–838. (doi:10.1038/nbt.2675)

100. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. & Church, G. M. 2013 RNA-guided human genome engineering via Cas9. *Science* **339**, 823–6. (doi:10.1126/science.1232033)

101. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. & Church, G. M. 2013 RNA-Guided Human Genome Engineering via Cas9. *Science (80-.).* **339**, 823–827. (doi:10.1126/science.1232033)

102. Till, B. J., Zerr, T., Comai, L. & Henikoff, S. 2006 A protocol for TILLING and Ecotilling in plants and animals. *Nat. Protoc.* **1**, 2465–2477. (doi:10.1038/nprot.2006.329)

103. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. & Rozen, S. G. 2012 Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40**, 1–12. (doi:10.1093/nar/gks596)

104. Koressaar, T. & Remm, M. 2007 Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**, 1289–1291. (doi:10.1093/bioinformatics/btm091)

105. Pruitt, K., Brown, G., Tatusova, T. & Maglott, D. 2002 The Reference Sequence (RefSeq) Database. *NCBI Handb.* , 1–24.

106. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J.

2009 Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191. (doi:10.1093/bioinformatics/btp033)

107. Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P. & Bork, P. 2000 SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* **28**, 231–234. (doi:gkd062 [pii])

108. Szklarczyk, D. et al. 2015 STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–52. (doi:10.1093/nar/gku1003)

109. Wu, C., MacLeod, I. & Su, A. I. 2013 BioGPS and MyGene.info: organizing online, gene-centric information. *Nucleic Acids Res.* **41**, D561–D565. (doi:10.1093/nar/gks1114)

110. Heng, T. S. P. & Painter, M. W. 2008 The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–4. (doi:10.1038/ni1008-1091)

111. Uhlen, M. et al. 2015 Tissue-based map of the human proteome. *Science (80-.).* **347**, 1260419–1260419. (doi:10.1126/science.1260419)

112. Eggington, J. M., Greene, T. & Bass, B. L. 2011 Predicting sites of ADAR editing in double-stranded RNA. *Nat. Commun.* **2**, 319. (doi:10.1038/ncomms1324)

113. Zhang, D. & Glotzer, M. 2014 Efficient site-specific editing of the C. elegans genome. *bioRxiv* , 007344. (doi:10.1101/007344)

114. Chromophore, R., Mishin, A. S., Subach, F. V, Yampolsky, I. V, King, W., Lukyanov, K. a & Verkhusha, V. V 2008 The First Mutant of the Aequorea V ictoria Green Fluorescent Protein That Forms a. , 4666–4673.

115. Yang, T. T. et al. 1998 Improved fluorescence and dual color detection with enhanced blue and green variants of the green fluorescent protein. *J. Biol. Chem.* **273**, 8212–8216. (doi:10.1074/jbc.273.14.8212)

116. Yamada, H. Y. & Rao, C. V. 2009 BRD8 is a potential chemosensitizing target for spindle poisons in colorectal cancer therapy. *Int. J. Oncol.* **35**, 1101–1109.

117. Doyon, Y., Selleck, W., Lane, W. S., Tan, S., Côté, J. & Co, J. 2004 Structural and Functional Conservation of the NuA4 Histone Acetyltransferase Complex from Yeast to Humans Structural and Functional Conservation of the NuA4 Histone Acetyltransferase Complex from Yeast to Humans. *Mol. Cell. Biol.* **24**, 1884–96. (doi:10.1128/MCB.24.5.1884)

118. Coletta, A., Pinney, J. W., Solís, D. Y. W., Marsh, J., Pettifer, S. R. & Attwood, T. K. 2010 Low-complexity regions within protein sequences have position-dependent roles. *BMC Syst. Biol.* **4**, 43. (doi:10.1186/1752-0509-4-43)

119. Wang, Y. et al. 2012 Coiled-coil networking shapes cell molecular machinery. *Mol. Biol. Cell* **23**, 3911–22. (doi:10.1091/mbc.E12-05-0396)

120. Cai, Y. et al. 2005 The mammalian YL1 protein is a shared subunit of the TRRAP/TIP60 histone acetyltransferase and SRCAP complexes. *J. Biol. Chem.* **280**, 13665–13670. (doi:10.1074/jbc.M500001200)

121. Jin, J. et al. 2005 A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. *J. Biol. Chem.* **280**, 41207–41212. (doi:10.1074/jbc.M509128200)

122. Rothenberg, M. E., Clarke, M. F. & Diehn, M. 2010 The Myc Connection: ES Cells and Cancer. *Cell* **143**, 184–186. (doi:10.1016/j.cell.2010.09.046)

123. Kim, J., Woo, A. J., Chu, J., Snow, J. W., Fujiwara, Y., Kim, C. G., Cantor, A. B. & Orkin, S. H. 2010 A Myc Network Accounts for Similarities between Embryonic Stem and Cancer Cell Transcription Programs. *Cell* **143**, 313–324. (doi:10.1016/j.cell.2010.09.010)

124. Orchard, S. et al. 2014 The MIntAct project--IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **42**, D358–63. (doi:10.1093/nar/gkt1115)

125. Chatr-Aryamontri, A. et al. 2015 The BioGRID interaction database: 2015 update. *Nucleic Acids Res.* **43**, D470–8. (doi:10.1093/nar/gku1204)

126. Victora, G. D. & Nussenzweig, M. C. 2012 Germinal Centers. *Annu. Rev. Immunol.* **30**, 429–457. (doi:10.1146/annurev-immunol-020711-075032)

127. Berek, C., Berger, a. & Apel, M. 1991 Maturation of the immune response in germinal centers. *Cell* **67**, 1121–1129. (doi:10.1016/0092-8674(91)90289-B)

128. Schwickert, T. A., Lindquist, R. L., Shakhar, G., Livshits, G., Skokos, D., Kosco-Vilbois, M. H., Dustin, M. L. & Nussenzweig, M. C. 2007 In vivo imaging of germinal centres reveals a dynamic open structure. *Nature* **446**, 83–87. (doi:10.1038/nature05573)

129. Allen, C. D. C., Ansel, K. M., Low, C., Lesley, R., Tamamura, H., Fujii, N. & Cyster, J. G. 2004 Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat. Immunol.* **5**, 943–952. (doi:10.1038/ni1100)

130. Zhou, P. et al. 2014 In vivo Discovery of Immunotherapy Targets in the Tumor Microenvironment. *Nature* **506**, 52–57. (doi:10.1161/STROKEAHA.113.002813.Oxidative)

131. Purwada, A., Jaiswal, M. K., Ahn, H., Nojima, T., Kitamura, D., Gaharwar, A. K., Cerchietti, L. & Singh, A. 2015 Ex vivo engineered immune organoids for controlled germinal center reactions. *Biomaterials* **63**, 24–34. (doi:10.1016/j.biomaterials.2015.06.002)

132. Maruyama, T., Dougan, S. K., Truttmann, M. C., Bilate, A. M., Ingram, J. R. & Ploegh, H. L. 2015 Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* , 1–8. (doi:10.1038/nbt.3190)

133. Heyn, H. & Esteller, M. 2015 An Adenine Code for DNA: A Second Life for N6-Methyladenine. *Cell* **161**, 710–713. (doi:10.1016/j.cell.2015.04.021)

134. Summerer, D. 2015 N6-Methyladenine: A Potential Epigenetic Mark in Eukaryotic Genomes. *Angew. Chemie Int. Ed.* **54**, 1014–10716. (doi:10.1002/anie.201504594)

135. Greer, E. L. et al. 2015 DNA Methylation on N6-Adenine in C. elegans. *Cell*

161, 1–11. (doi:10.1016/j.cell.2015.04.005)

136. Yue, Y., Liu, J. & He, C. 2015 RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes Dev.* **29**, 1343–1355. (doi:10.1101/gad.262766.115.GENES)

137. Garabedian, M. V. et al. 2012 The double-bromodomain proteins Bdf1 and Bdf2 modulate chromatin structure to regulate S-phase stress response in schizosaccharomyces pombe. *Genetics* **190**, 487–500. (doi:10.1534/genetics.111.135459)

138. Jackson, S. P. & Bartek, J. 2009 The DNA-damage response in human biology and disease. *Nature* **461**, 1071–1078. (doi:10.1038/nature08467)

139. Kang, H. J., Le, T. V. T., Kim, K., Hur, J., Kim, K. K. & Park, H. J. 2014 Novel interaction of the Z-DNA binding domain of human ADAR1 with the oncogenic c-myc promoter G-quadruplex. *J. Mol. Biol.* **426**, 2594–2604. (doi:10.1016/j.jmb.2014.05.001)

140. Addgene In press. https://www.addgene.org/crispr/reference/history/.

12 Supplemental figures

12.1 Assessment of gRNAs and wild-type Cas9 plasmids

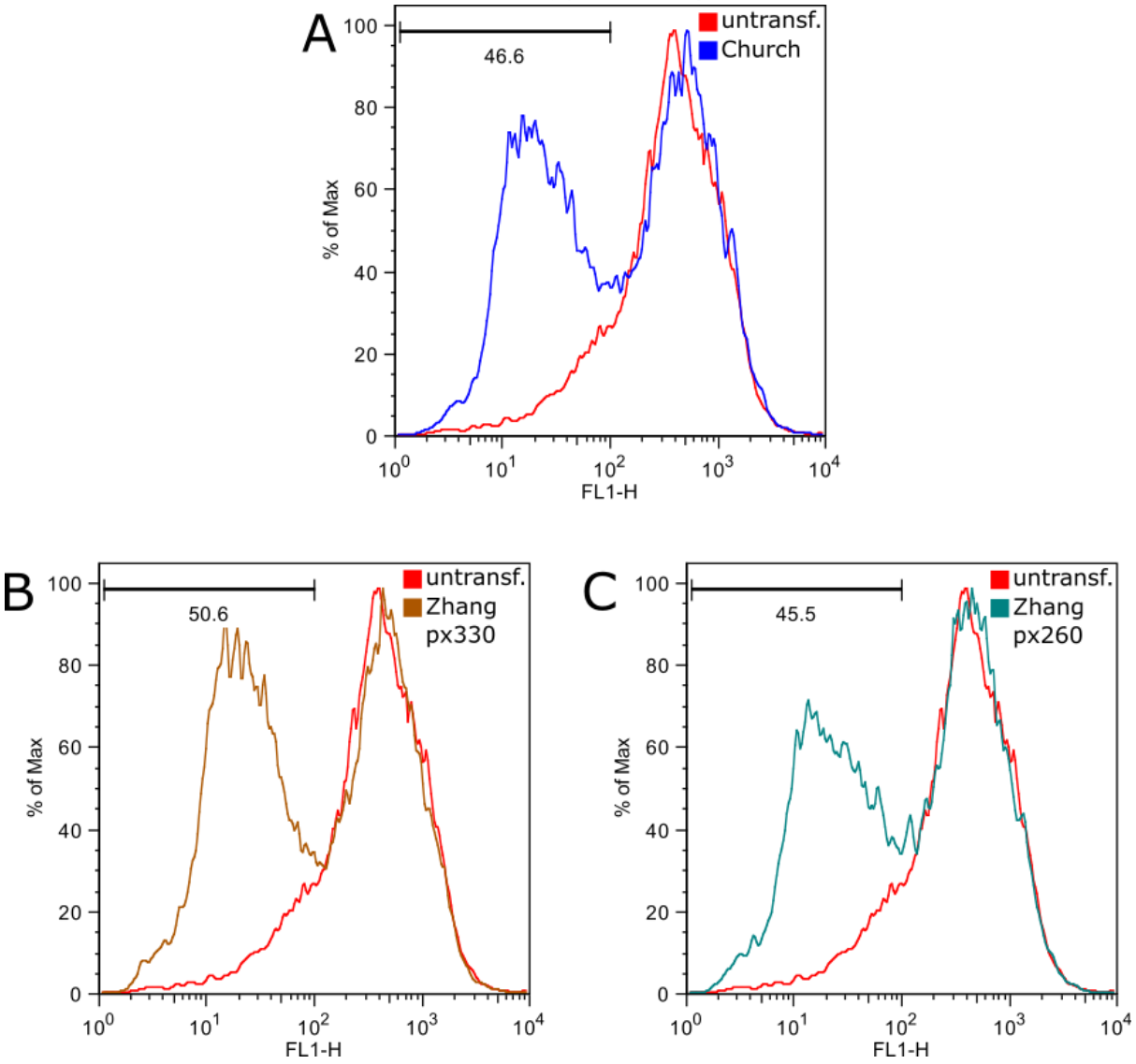

Figure S 1: FACS results for wildtype Cas9 transfections of 293T/GFP-puro cells.
One Church plasmid and two Zhang plasmids were tested and compared to an untransfected control.

12.2 Human BRD8 domains and conservation

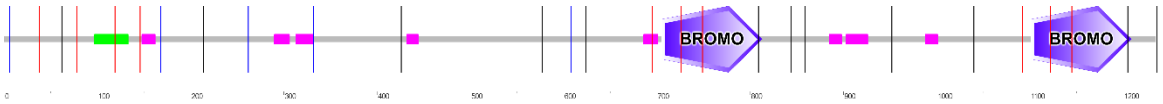

Figure S 2: Human long isoform BRD8 domain prediction by SMART. Green = coiled-coil. Pink = low-complexity region.

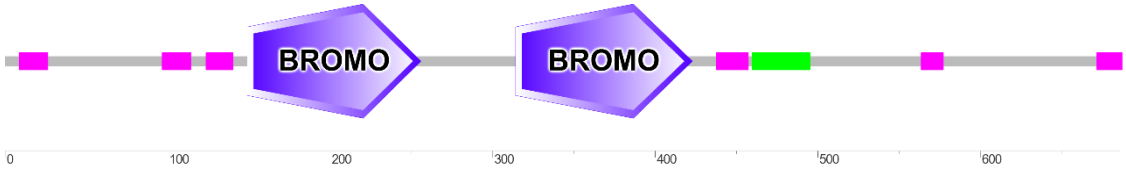

Figure S 3: S-cerevisiae Bdf1 domain prediction by SMART. Green = coiled-coil. Pink = low-complexity region.

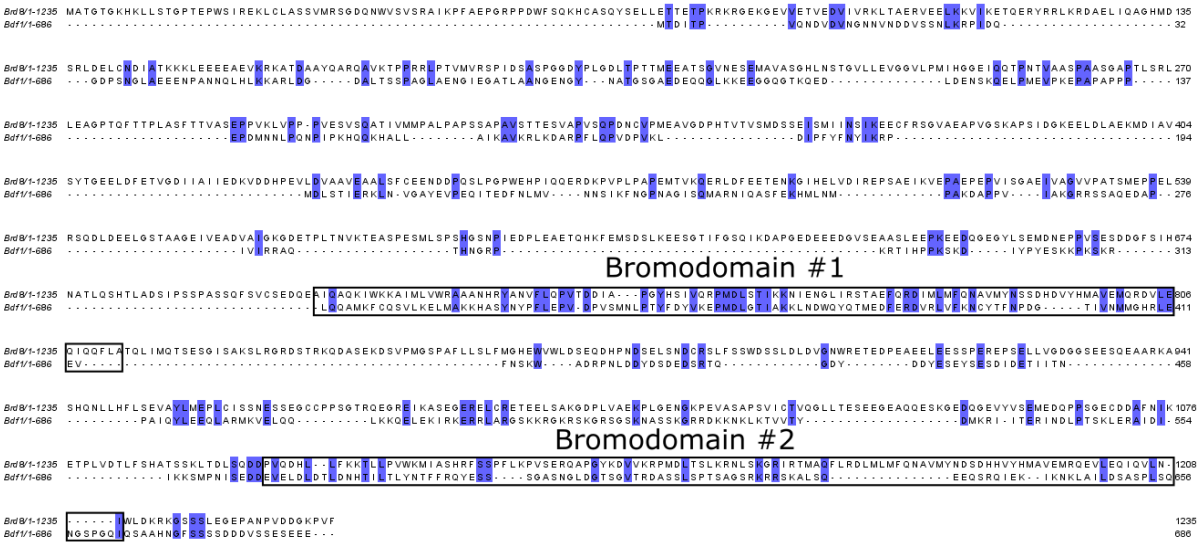

Figure S 4: Pairwise alignment of human long isoform BRD8 and S.cerevisiae Bdf1

Figure S 5: Multiple alignment of the long isoform BRD8 from Homo sapiens, Pan troglodytes, Cricetulus griseus, Sorex araneus, Condylura cristata, Felis catus, Ovis aries, Equus przewalskii, Pteropus vampyrus, Odobenus rosmarus divergens, Tursiops truncates, Alligator sinensis and Chelonia mydas.